PRIMATES DEL MIOCENO • RECONSIDERACION DEL CEREBELO

INVESTIGACION y CIENCIA

OCTUBRE 2003
5,50 EURO

Edición española de SCIENTIFIC AMERICAN

LA INFORMACION EN EL UNIVERSO HOLOGRAFICO

INTERFERENCIA DE ARN

INFORMATIZACION Y DESARROLLO HUMANO

EL ORACULO DE DELFOS

DINAMICA DE AGREGADOS ATOMICOS

LAS ESTRELLAS BINARIAS



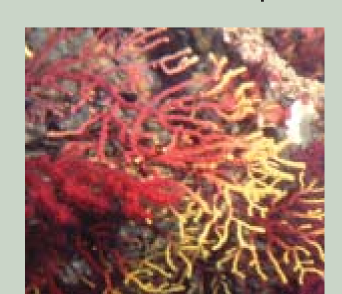


Edición española de SCIENTIFIC AMERICAN

Octubre de 2003

Número 325











DIRECTOR GENERAL José M.ª Valderas Gallardo
DIRECTORA FINANCIERA Pilar Bronchal Garfella
EDICIONES Juan Pedro Campos Gómez
PRODUCCIÓN M.ª Cruz Iglesias Capón
Bernat Peso Infante
SECRETARÍA Purificación Mayoral Martínez
ADMINISTRACIÓN Victoria Andrés Laiglesia
SUSCRIPCIONES Concepción Orenes Delgado
Olga Blanco Romero
EDITA Prensa Científica, S. A. Muntaner, 339 pral. 1.ª
08021 Barcelona (España)
Teléfono 934 143 344 Telefax 934 145 413
www.investigacionyciencia.es

SCIENTIFIC AMERICAN

EDITOR IN CHIEF John Rennie
EXECUTIVE EDITOR Mariette DiChristina
MANAGING EDITOR Ricki L. Rusting
NEWS EDITOR Philip M. Yam
SPECIAL PROJECTS EDITOR Gary Stix
REVIEWS EDITOR Michelle Press
SENIOR WRITER W. Wayt Gibbs
EDITORS Mark Alpert, Steven Ashley, Graham P. Collins, Carol Ezzell, Steve Mirsky y George Musser
PRODUCTION EDITOR Richard Hunt
VICE PRESIDENT AND MANAGING DIRECTOR, INTERNATIONAL Dean Sanderson
PRESIDENT AND CHIEF EXECUTIVE OFFICER Gretchen G. Teichgraeber
CHAIRMAN Rolf Grisebach

DISTRIBUCION

para España:

LOGISTA, S. A.
Aragoneses, 18
(Pol. Ind. Alcobendas)
28108 Alcobendas (Madrid)
Tel. 914 843 900

para los restantes países:

Prensa Científica, S. A.
Muntaner, 339 pral. 1.ª
08021 Barcelona
Teléfono 934 143 344

PUBLICIDAD

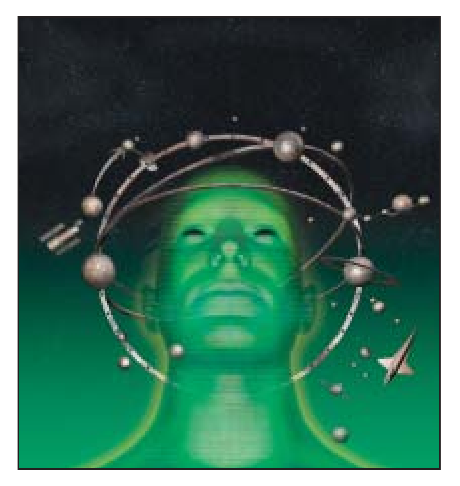
GM Publicidad
Edificio Eurobuilding
Juan Ramón Jiménez, 8, 1.ª planta
28036 Madrid
Tel. 912 776 400
Fax 914 097 046

Cataluña:
QUERALTO COMUNICACION
Julián Queraltó
Sant Antoni M.ª Claret, 281 4.º 3.ª
08041 Barcelona
Tel. y fax 933 524 532
Móvil 629 555 703

COLABORADORES DE ESTE NUMERO

Asesoramiento y traducción:

Esteban Santiago: *Interferencia de ARN*; Luis Bou: *Informatización y desarrollo humano, Reconsideración del cerebelo* y *Aventuras problemáticas*; Ramón Pascual: *La información en el universo holográfico*; José M. García de la Mora: *El oráculo de Delfos*; Julio A. Alonso: *Dinámica de agregados atómicos*; Carlos Lorenzo: *Primates del Mioceno*; J. Vilardell: *Hace...*, *Apuntes* e *Ideas aplicadas*; Jürgen Goicoechea: *Curiosidades de la física*



Portada: Kenn Brown

SUSCRIPCIONES

Prensa Científica S. A.
Muntaner, 339 pral. 1.ª
08021 Barcelona (España)
Teléfono 934 143 344
Fax 934 145 413

Precios de suscripción:

	Un año	Dos años
España	60,00 euro	110,00 euro
Extranjero	85,00 euro	160,00 euro

Ejemplares sueltos:

Ordinario: 5,50 euro
Extraordinario: 6,00 euro

—El precio de los ejemplares atrasados es el mismo que el de los actuales.



ISSN 0210136X Dep. legal: B. 38.999 – 76

Imprime Rotocayfo-Quebecor, S.A. Ctra. de Caldes, km 3 - 08130 Santa Perpètua de Mogoda (Barcelona)

Printed in Spain - Impreso en España



HACE

...cincuenta años

VERDURAS PARA CENAR. «Numerosos científicos de todo el mundo muestran su interés por las posibilidades alimenticias de unas plantas acuáticas, las algas. Se deduce de experimentos de laboratorio que cada hectárea dedicada al cultivo del alga *Chlorella* produciría al año alrededor de media tonelada de proteína y unos 75 kilogramos de grasa, cifras astronómicas comparadas con las tasas de producción actuales de la agricultura clásica. Que las algas puedan contribuir de modo apreciable al abastecimiento mundial de alimentos dependerá del costo y del rendimiento de los cultivos a gran escala. La producción de una tonelada de proteína de alga requiere del orden de 1,1 toneladas de nitrato potásico y 0,75 toneladas de sulfato amónico.»

...cien años

EL FRACASO DE LANGLEY. «A quienes de veras se interesan por la navegación aérea les pesará el fracaso del último experimento del profesor Samuel Pierpoint Langley, no tanto porque el aparato se negase a volar, sino por los adversos comentarios que el intento ha suscitado en la prensa. El artefacto es el resultado de años de arduos estudios y experimentación incesante. Que haya fallado debe considerarse simplemente como una etapa más hacia la solución del problema de la navegación aérea, y no como un rotundo fracaso sin salvación posible. En el informe del profesor C. L. Manley consta que se halló dañado el embrague que sujetaba el aparato en la guía de lanzamiento; debió de soltarse en el instante de la caída.» [*Nota de la redacción: el fracaso de esta prueba y el de otra posterior, del 8 de diciembre de 1903, produjeron unas críticas públicas tan cáusticas que Langley renunció a sus investigaciones en el campo de la aviación.*]

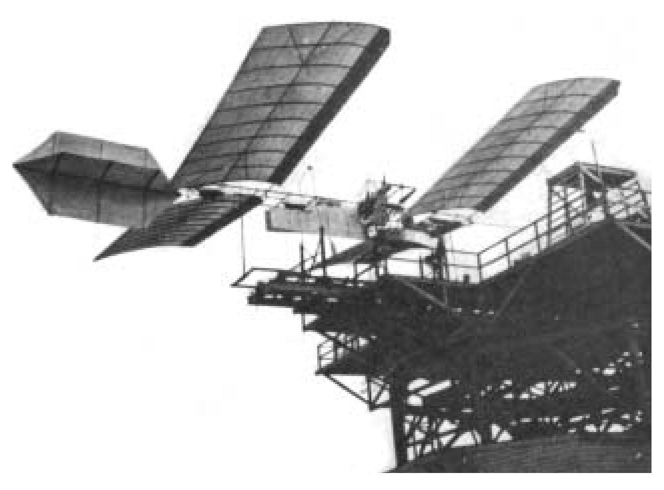
¿ME DA FUEGO? «Por ley del 10 de mayo de 1903, Alemania ha prohibido el uso de fósforo blanco en la fabricación de cerillas. El gobierno ha comprado una nueva sustancia, hecha de fósforo rojo y clorato potásico, que sustituirá al deletéreo y peligroso fósforo blanco. Pese a su elevado punto de ignición, la nueva sustancia puede encenderse frotándola contra casi cualquier material —lija, ladrillo, suela de zapato, tela áspera, etcétera—. Es una gran ventaja que no se encienda fácilmente, e importante: basta recordar los incendios que ha causado la ignición de las cerillas de fósforo blanco por acción de los rayos solares.»

SALUBRIDAD AMERICANA. «Desde la ocupación americana de Cuba, la fiebre amarilla se está gradualmente erradicando. Tan notable cambio sanitario se debe en parte a la destrucción por los médicos militares de las viejas supersticiones, en parte al exterminio sistemático de los mosquitos. Los expertos de nuestro ejército han disipado la creencia, tan común, de que la fiebre amarilla es una enfermedad mortal terriblemente contagiosa; han demostrado que la transmiten los mosquitos. Hasta tal punto han sido fructíferas las tentativas de exterminar el insecto que, con el tiempo, las ciudades cubanas estarán tan libres de la fiebre amarilla como nuestros puertos del sur.»

HORMIGAS. «Una consecuencia imprevista del avenamiento de Nueva Orleans es la aparición de hordas de hormigas, que se han vuelto tan amenazadoras como las plagas de Egipto. Atacan la obra de carpintería de las casas. Rápidamente la destruyen y se abren camino hacia los almacenes donde se guardan productos valiosos. En terrenos húmedos no pueden reproducirse; pero ahora que ya no hay humedad, se resisten a la extinción.»

...ciento cincuenta años

CARBÓN PARA LOS TRENES. «Con pocas excepciones, la madera es el único combustible de los motores de locomotora. Está escaseando y encareciéndose tanto que debe buscársele un sustituto. En primer lugar, se propone la antracita, el más barato y el que menos humos y residuos produce. Pero se cree que destruye con tal rapidez la caldera que su uso es ineconómico. Otras objeciones se basan en la intensidad del calor. Sin embargo, el motor Millholland obvia esos inconvenientes. La compañía Ferrocarriles de Reading, en Pennsylvania, utiliza ya para su servicio regular veintiocho locomotoras de primera clase, basadas en el sistema Millholland, que emplean exclusivamente antracita. No hay ingeniero que recurra a una locomotora que queme madera si puede conseguir una de carbón; dan mucho menos trabajo y producen un mayor rendimiento.»



El avión de Langley en su catapulta, 1903

APUNTES

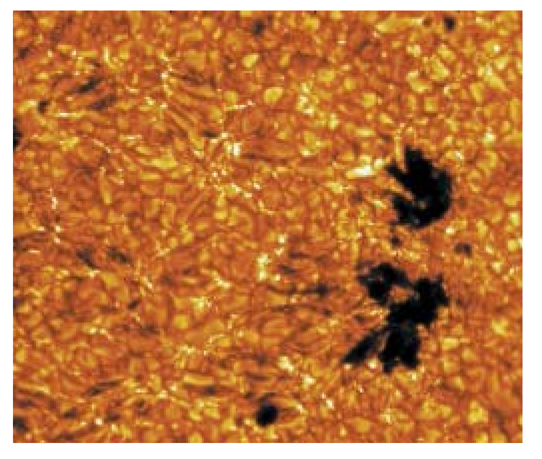
FISICA DE FLUIDOS

Un martillo de agua luminoso

La sonoluminiscencia consiste en la generación de luz en un fluido por medio de ondas sonoras. A finales de 2002 el fenómeno dio lugar a una polémica. La revista *Science* publicó un artículo que anunciaba la consecución de la fusión del hidrógeno por medio de la sonoluminiscencia (en la que se genera en un punto un calor muy intenso). Muchos lo compararon despectivamente a la fusión fría, el espejismo científico que generó tanto revuelo años atrás, y acusaron a la revista de publicar un material poco de fiar. Pero la sonolumiscencia en sí es respetable. Hace ocho años publicamos un artículo de Seth Putterman donde se abordaban sus fundamentos, aún no del todo claros. Ahora, Putterman y sus colaboradores acaban de dar a conocer una nueva forma de producirla, más eficaz que enviar ondas sonoras hacia el interior del fluido, como se venía haciendo. Un tubo de unos 50 centímetros de largo y unos 4 de diámetro, relleno de agua con un poco de xenón diluido, un "martillo de agua", se agita, con una fuerza que dobla la de la gravedad, de manera que el líquido se mueva hacia un extremo distinto en cada mitad del tubo y se abra un vacío de un centímetro en medio. Cuando el vacío se cierra, una onda de choque crea burbujas en el agua; su derrumbe emite impulsos de luz, en su mayor parte ultravioletas, de menos de un nanosegundo de duración.

ASTRONOMIA

Nueva luz sobre el viejo Sol



Un equipo encabezado por Tom Berger, del Laboratorio de Astrofísica y del Sol de Lockheed Martin, en Palo Alto (California), ha conseguido las fotos de mayor resolución que se hayan tomado del Sol. Las imágenes, donde se distinguen detalles de sólo 75 kilómetros de ancho, revelan una fotosfera, antaño supuesta plana y sin rasgos, dotada de una sorprendente estructura. Los gránulos, cada uno mayor que España, resultan del borboteo del calor procedente del interior; las manchas y demás "poros" oscuros se muestran hundidos en la superficie. Las fáculas (zonas de mayor brillo entre los gránulos) parecen elevarse por encima de la superficie; quizá sean la razón del aumento de actividad durante los máximos solares.

—*Philip Yam*

La superficie del Sol aparece plagada de gránulos y otras estructuras

TOM BERGER *Laboratorio de Astrofísica y del Sol Lockheed Martin*

HISTORIA DE LA FISICA

La fragilidad de la memoria

La investigación histórica puede demorarse, invertir esfuerzos considerables, en asuntos que parecen minucias. Un doctorando italiano, Alberto De Gregorio, ha encontrado pruebas documentales de que el equipo de Fermi no descubrió los efectos de la moderación de neutrones, que serían fundamentales para el control de la fisión nuclear, el 22 de octubre de 1934, sino dos días antes. De Gregorio percibe en este hallazgo la importancia de un síntoma. La historia de aquellos días se ha escrito a partir de los recuerdos, contados casi medio siglo después, de algunos miembros de aquel célebre grupo, como Segrè, Pontecorvo —que se pasaría a la Unión Soviética, famoso episodio de la guerra fría, donde enunciaría la hipótesis de las oscilaciones neutrínicas— y Amaldi. De Gregorio piensa que la fecha que aportaron procede en realidad de la biografía de Fermi escrita por su viuda en 1954. El error se habría propagado desde ahí. ¿No les fallaría también la memoria en más cosas? Segrè, por ejemplo, explicó que Fermi no quería poner nombre a los transuránidos que creía haber descubierto, cuando consta que en su laboratorio los llamaban ausonium y hesperium. La disolución del pasado con el tiempo deja también lugar para el misterio. Los documentos consultados por De Gregorio se encuentran en la *Domus Galilaeana* de Pisa, que alberga un vasto archivo de los trabajos de Fermi en su etapa italiana. La colección acaba el 26 de octubre de 1938. Ese mismo día desapareció, sin que nunca más se volviera a saber de él, otro de los grandes físicos italianos, Ettore Majorana. Leonardo Sciascia escribió un libro sobre éste, *La scomparsa de Majorana,* donde se permitía imaginar, a partir de un vago y lejano rumor, si no se habría retirado a la cartuja en la que quizá se recluyó también un tripulante del *Enola Gay*. Pero, pese a que fuera autor de novelas entre filosóficas y policiacas, no soñó con el extraño nexo que insinúa De Gregorio.

MICROMECANICA

Motores bacterianos

La afición por combinar microbios y máquinas se extiende entre los ingenieros de sistemas micromecánicos. Con paciencia, algunos han logrado que los rotativos flagelos de ciertas bacterias actúen como bombas y válvulas. Ahora hay investigadores que uncen microbios para que eleven y muevan objetos de una manera que recuerda mucho a los motores fuera borda. *Serratia marcescens* se adhiere por sí sola a las superficies y facilita así que se las enganche a otros dispositivos, según Linda Turner, del Instituto Rowland, en la Universidad de Harvard. Hasta unas 50 pueden cubrir una cuenta de plástico del tamaño de una célula sanguínea; cuando están muy compactadas, los flagelos se influencian entre sí y la coordinación mejora. Turner espera guiar las bacterias, que nadan más o menos a un milímetro por minuto, con señales luminosas o químicas. Un tapiz de bacterias repartiría las sustancias cargadas de gas con mayor rapidez que la mera difusión, y facilitaría la mezcla y agitación de fluidos melosos.

—Charles Choi

25 micras

LINDA TURNER *Instituto Roland*

Transportan esta placa de silicona unas bacterias que tiene debajo. La llevan de derecha a izquierda; se ha fotografiado el movimiento a intervalos de cinco segundos

PERCEPCION

Golpe por golpe

Cuando dos niños que se pelean se quejan de que el otro ha golpeado más fuerte, puede que ambos digan la verdad. Investigadores del Colegio Universidad de Londres realizaron experimentos en los que a parejas de sujetos se les decía que devolvieran, con la misma fuerza, los golpes que recibieran en un dedo. La violencia se intensificaba rápidamente: los voluntarios aumentaban su fuerza en un 38 por ciento a cada turno. Conjeturan esos científicos que los sujetos se quedaban cortos al calcular la fuerza que aplicaban porque el cerebro, cuando ha de planear un movimiento, puede que atenúe la sensación que producirá. Quizá libere así recursos neurales que dejen al cerebro mejor dispuesto para recibir estímulos. Con el objeto de respaldar su teoría, se pidió a los participantes que devolvieran los golpes con un *joystick*, no directamente con los dedos. Ese método esquivó los mecanismos predictivos del cerebro: los sujetos reproducían exactamente la violencia sufrida.

—Philip Yam



RUBBERBALL *Getty*

Calibrar exactamente la fuerza de un golpe depende de que estemos dando o recibiendo

BIOFISICA

Vulnerabilidad del ADN

Es bien sabido que los electrones con gran energía —del orden de megaelectrón-volt, alrededor de un millón de veces la de un fotón de luz visible— dañan el ADN. No es su impacto directo en los átomos la causa de ese deterioro, sino la estela de electrones de menor energía, entre 1 y 20 electrón-volt (eV), que dejan a su paso. Antes de 2000, se pensaba que un electrón secundario había de tener más de 10 eV de energía para degradar el ADN —la necesaria para ionizarlo—. Pero un grupo de Quebec demostró que bastaban 3 eV. Parece que estos electrones débiles atacan ambas hebras de la macromolécula ligándose a una molécula de una de ellas. El ion resultante se desintegra; los fragmentos dañan la otra hebra por medio de reacciones químicas. La maquinaria reparadora del ADN puede restañar lesiones sueltas, pero fallará cuando se acumulen en poco espacio o se produzcan alteraciones complejas. El grupo de Tilmann Märk, de Innsbruck, ha observado ahora la nocividad de electrones de bastante menos de 1 eV (una energía propia de luz infrarroja). No han estudiado moléculas de ADN en sí, sino nubes gaseosas de bases de ADN, portadoras de información, y de dexosirribosa, uno de los componentes estructurales. También ahí parece que las moléculas se fragmentan tras incorporarse los electrones. Han comprobado además que la sustitución de un hidrógeno por cloro en las bases uracilo centuplica la probabilidad de que los electrones las desintegren. Los de Quebec, con moléculas enteras, estudian la sustitución por bromo. La sensibilización depende de la secuencia de bases en que se introduzca la sustitución; podría indicar una manera de atacar las células tumorales en sitios concretos del ADN.

—Graham P. Collins

El uranio emite partículas alfa; cada una genera en un tejido 160.000 electrones de baja energía

DEPARTAMENTO DE ENERGIA DE EE.UU. *SPL/Photo Researchers, Inc.*



Interferencia de ARN

La mayoría de las células animales y vegetales contienen un sistema de silenciamiento de genes. Se valen del mismo para triturar el ARN que producen

Nelson C. Lau y David P. Bartel

Cuando se observa una célula viva sobre el porta de un microscopio parece sosegada. Pero debajo de esa quietud externa se esconde una frenética actividad bioquímica. El genoma de ADN del interior de una célula vegetal o animal contiene millares de genes. Dejada a su aire, la maquinaria de transcripción de la célula expresaría de inmediato cada uno de los genes del genoma: desenrollaría la doble hélice del ADN, transcribiría cada gen en su correspondiente ARN mensajero de una hebra (ARNm) y, por último, traduciría los mensajes de ARN en sus proteínas.

Ninguna célula podría trabajar con semejante alboroto. Deben las células silenciar la mayoría de sus genes, para permitir que entre en acción el subgrupo apropiado. En la mayoría de los casos, el código del ADN de un gen se transcribe en ARNm sólo si una determinada proteína se ha acoplado, en su ensamblaje, a una región reguladora especial del gen.



JANA BRENNING

Algunos genes, sin embargo, son tan subversivos, que nunca se les debería conceder la libertad de expresión. Si a los genes de los elementos genéticos móviles se les dejara emitir sus mensajes de ARN, podrían saltar de un punto a otro del ADN produciendo cáncer u otras enfermedades. En la misma senda, si se permitiera a los virus expresar su mensaje sin ningún control, no tardarían en secuestrar los mecanismos celulares de síntesis de proteínas para aplicarlos a la fabricación de proteínas víricas.

Las células disponen de medios preventivos y de contragolpe. Hace ya algún tiempo se identificó un sistema, la respuesta de interferón, que las células humanas despliegan cuando los genes víricos penetran en su interior. La respuesta puede impedir la expresión de casi todos los genes. El proceso recuerda la parada de las prensas por orden gubernativa.

En fecha más reciente se descubrió un mecanismo de seguridad mucho más enérgico y preciso; lo poseen casi todas las células animales y vegetales. A este sistema censor se le conoce por interferencia de ARN, abreviado en iARN. Cuando se expresa un gen amenazador, la maquinaria de la iARN corre a silenciarlo: intercepta y destruye sólo el ARNm transgresor, sin perturbar los mensajes de los genes restantes.

Conforme se ha ido sondeando el modo de operar del censor celular y determinando los estímulos que lo activan, ha aumentado el interés de los expertos. Pues, en línea de principio, cabe la posibilidad de idear formas de obligar a la iARN a silenciar genes implicados en el cáncer, en las infecciones víricas y en otras patologías. De lograr materializarlo, nos hallaríamos ante un prometedor campo, inédito, para la medicación.

Mientras llega ese día, los investigadores que trabajan con plantas, moscas y otros organismos experimentales han conseguido ya que la iARN suprima casi cualquiera de los genes en estudio, lo que les permite intuir la función del gen. En cuanto herramienta de investigación, la iARN ha tenido un éxito inmediato. Merced a ella, centenares de laboratorios abordan ya cuestiones que estaban lejos de su alcance hace escasos años.

Aunque la mayoría de los grupos de investigación utilizan la interferencia de ARN como medio para conseguir un fin, algunos se han centrado en la propia naturaleza del fenómeno. A nosotros nos interesa el papel de la maquinaria de la iARN en el desarrollo normal de plantas, hongos y animales, sin excluir al hombre.

Un silencio extraño

Las primeras pistas sobre la existencia de la iARN aparecieron trece años atrás. Richard A. Jorgensen, hoy en la Universidad de Arizona, y Joseph Mol, de la Libre de Amsterdam, cada uno por su cuenta, insertaron en petunias de flor púrpura copias adicionales del gen de su pigmento nativo. Confiaban en que las plantas, así manipuladas, dieran flores de un violeta más intenso. Pero los pétalos de sus petunias se tiñeron de manchas blancas.

Jorgensen y Mol dedujeron que las copias extra habían despertado la censura de los genes del pigmento púrpura, incluidos los genes naturales de las petunias; el fenómeno provocó la aparición de flores variegadas o albinas. A esta censura doble, sobre un gen insertado y el gen nativo correspondiente, la llamamos cosupresión. Más tarde se observaría en hongos, mosca del vinagre y otros organismos.

Las primeras pistas para desentrañar el misterio en torno al silenciamiento de los genes llegaron unos años después. William G. Dougherty, de la Universidad estatal de Oregón, había empezado con su grupo a investigar plantas de tabaco especiales: habían sido manipuladas genéticamente para que incluyeran en su ADN varias copias del gen CP del virus del grabado del tabaco (*Marmor erodens*). El gen CP determina la proteína de la cubierta vírica.

Cuando estas plantas se expusieron al virus, algunas permanecieron inmunes a la infección. Dougherty atribuyó la inmunidad a la cosupresión. Las plantas reaccionaban ante la expresión inicial de sus genes CP de origen foráneo mediante la inactivación de dicha expresión y, luego, mediante el bloqueo también de la





1. LAS PRIMERAS PISTAS sobre la existencia de censores de genes en el reino vegetal se recabaron entre las petunias púrpura. Tras insertar genes adicionales de pigmentos en plantas normales (*izquierda*), echaron flores ribeteadas de blanco o con extensas zonas albinas (*centro* y *derecha*).

RICHARD A. JORGENSEN *Universidad de Arizona*

expresión del gen CP del virus invasor (que necesita la proteína de la cubierta para producir la infección). El laboratorio de Dougherty prosiguió su trabajo. Demostró que la inmunidad no requería la síntesis de la proteína de la cubierta por las plantas; debía haber algo, relacionado con el ARN transcrito desde el gen CP, que diera cuenta de la resistencia opuesta por las plantas a la infección.

El grupo reveló que las plantas no sólo podrían inactivar genes específicos en los virus, sino que también estos agentes podrían disparar el silenciamiento de genes seleccionados. Algunas plantas de Dougherty que no lograron suprimir por sí mismas sus genes CP, se infectaron con el virus, que se replicaba sin problemas en las células vegetales. Cuando los investigadores midieron, más adelante, el ARN transcrito por los genes CP de las plantas afectadas, vieron que estos mensajes se habían desvanecido casi por completo: la infección había provocado la inactivación de los genes CP.

Resumen/*Interferencia de ARN*

- Se ha convertido en práctica rutinaria la inserción, en organismos experimentales, de genes manipulados. Se trata de una técnica conocida desde hace cierto tiempo. Muy reciente, sin embargo, es el descubrimiento de un medio eficaz y apropiado para silenciar, en el mismo interior celular, un gen específico.
- En su mayoría, las células animales y vegetales poseen mecanismos internos que emplean formas insólitas de ARN, molécula mensajera genética, para acallar determinados genes de un modo natural.
- Esta maquinaria ha evolucionado. Protege de genes hostiles a la célula y regula la actividad de los normales durante el desarrollo celular. La investigación del proceso abre el camino para la fabricación de fármacos que se apoyen en los mecanismos de interferencia de ARN para prevenir o curar ciertas patologías.

Mientras tanto los biólogos que experimentaban con el nemátodo *Caenorhabditis elegans*, un gusano diminuto y transparente, andaban desconcertados en su empeño por averiguar qué ocurría cuando usaban ARN "antisentido" para inactivar los genes que estudiaban. Propio del ARN antisentido es emparejarse con una secuencia determinada de ARNm, a la manera en que dos hebras complementarias de ADN se entrelazan para formar una hélice doble. Cada hebra de ADN o de ARN es una cadena de nucleótidos, representados por las letras A, C, G y U (este último en el ARN) o T (en el ADN). Un nucleótido de C se enlaza siempre con otro de G, y uno de A se empareja con otro de U o de T. Una hebra de ARN antisentido se une con la hebra de un ARNm complementario y crean una estructura bicatenaria, incapaz de traducirse en una proteína útil.

Con el paso de los años, estos experimentos con ARN antisentido acometidos en diversos organismos alcanzaron un éxito desigual. Sin embargo, fue para todos una sorpresa que el ARN "con sentido" bloqueara la expresión del gen. El ARN con sentido tiene la misma secuencia que el ARNm diana, razón por la cual no puede formar doble hélice con el ARN mensajero.

La escena estaba preparada para el experimento decisivo. Se llevó a cabo, unos cinco años más tarde, en los laboratorios de Andrew Z. Fire, de la Institución Carnegie de Washington, y Craig C. Mello, de la facultad de medicina de la Universidad de Massachusetts. Fire y Mello sospecharon que las preparaciones anteriores de ARN con sentido y ARN antisentido que se inyectaban a los gusanos carecían de la pureza exigida; probablemente, contenían trazas de ARN bicatenario. Quizás ese ARN bicatenario alertaba a los censores, pensaron.

Para comprobar sus sospechas, Fire, Mello y sus colaboradores inocularon en los nemátodos ora ARN monocatenario, ora ARN bicatenario, correspondiente al gen *unc-22*. (Este gen desempeña un papel destacado en la función muscular.) Las cantidades elevadas de ARN *unc-22* monocatenario, con sentido o antisentido, apenas si ejercían efecto alguno en los gusanos. Pero bastaban unas pocas moléculas del ARN *unc-22* bicatenario para inducirles espasmos, en ellos y en su progenie. Se hallaban ante un signo inequívoco de que algo había comenzado a interponerse en el camino de la expresión del gen *unc-22*. Fire y

Mello observaron idéntico, y sorprendente, efecto silenciador en casi todos los genes que estudiaron, desde los musculares a los genes de la fertilidad y viabilidad. Denominaron a este fenómeno "interferencia de ARN" (iARN), para resaltar el papel clave del ARN bicatenario en la censura del gen correspondiente.

Los estudiosos de vegetales y hongos centraron también su atención sobre el ARN bicatenario en su búsqueda del responsable del silenciamiento. Demostraron que las hebras de ARN que podían replegarse sobre sí mismas para formar largas tiras de ARN bicatenario constituían potentes inductores del silenciamiento. Otros análisis revelaron que, en la cosupresión, se requería un gen que capacita a las células para convertir ARN monocatenario en ARN bicatenario.

De tal gavilla de hallazgos cabía inferir que las petunias de Jorgensen y Mol reconocían los genes extra del pigmento como insólitos (por un mecanismo que sigue siendo misterioso) y convertían su ARN mensajero en ARN bicatenario, que provocaba el silenciamiento tanto de los genes extra como de los nativos.

El concepto de un ARN bicatenario inductor explica también por qué la infección vírica amordazaba los genes CP en las plantas de Dougherty. El virus del tabaco *Marmor erodens* había producido ARN bicatenario del genoma vírico completo a medida que iba reproduciéndose, un fenómeno que acontece con muchos virus. Las células vegetales respondían bloqueando los mensajes de ARN de todos los genes asociados con el virus, incluidos los genes CP incorporados en el ADN de la planta.

Nadie entendía que semejante sistema, enérgico y ubicuo, de regulación de la expresión génica hubiera pasado inadvertido tanto tiempo. Levantado el velo del misterio que le envolvía, genéticos y bioquímicos se han volcado ahora en el análisis de su mecanismo de acción.

Corte y empalme de los mensajes genéticos

Pronto se confirmó el fenómeno de la interferencia de ARN en algas, platelmintos y mosca del vinagre; ramas dispares del árbol evolutivo. La demostración de su presencia en células humanas típicas y de otros mamíferos resultó, sin embargo, una empresa harto más compleja.

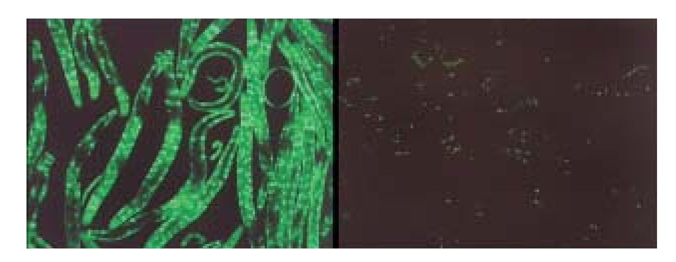
Cuando una célula humana se infecta con un virus que produce ARN bicatenarios, puede entrar en un estado que podríamos llamar de "cerrojazo": la enzima PKR bloquea la traducción de todos los ARN mensajeros —normales y víricos— y otra enzima, la ARNasa L, destruye indiscriminadamente los ARN mensajeros. Estas reacciones ante los ARN bicatenarios se integran en la respuesta al interferón; se les considera componentes de la misma porque se disparan más fácilmente después de que las células hayan estado expuestas a interferones, moléculas que las células infectadas segregan para alertar de peligro a las células vecinas.

Por desgracia, cuando los investigadores introducen ARN bicatenarios artificiales (los empleados para inducir la interferencia de ARN en gusanos y moscas) en el interior de células de mamíferos maduros, la respuesta al interferón inactiva todos y cada uno de los genes de la célula. Se requería, pues, conocer mucho mejor el funcionamiento de la interferencia de ARN antes de que pudiera usarse de manera rutinaria sin hacer saltar las alarmas del interferón. Además de los investigadores pioneros, que ya hemos mencionado, Thomas Tuschl, de la Universidad Rockefeller, Philip D. Zamore, de la facultad de medicina de la Universidad de Massachusetts, Gregory Hannon, del Laboratorio Cold Spring Harbor de Nueva York, y muchos otros han contribuido a lo que hoy sabemos sobre el mecanismo de la transferencia de ARN.

¿Cómo opera la iARN? En el interior de una célula, el ARN bicatenario sale al encuentro de la enzima "Dicer" ("la tajadora"). Mediante el proceso químico de hidrólisis, Dicer trocea las cadenas largas de ARN en fragmentos, a los que se da el nombre de ARN de interferencia pequeños (ARNip). Cada ARNip consta de unos 22 nucleótidos.

La enzima Dicer corta el ARN bicatenario por puntos especiales, en lugares ligeramente no coincidentes; por ello, cada ARNip resultante tiene dos nucleótidos sin emparejar en un extremo. El dúplex ARNip se desenrolla entonces; una de las cadenas del dúplex se carga en un ensamblaje de proteínas para formar el complejo silenciador inducido por ARN (CSIR).

Dentro del complejo silenciador, la molécula de ARNip se acomoda de suerte tal, que las moléculas de ARN mensajero puedan encontrarse con ella. El CSIR se las verá con millares de ARNm diferentes que hay en cualquier célula en cualquier momento. Pero el

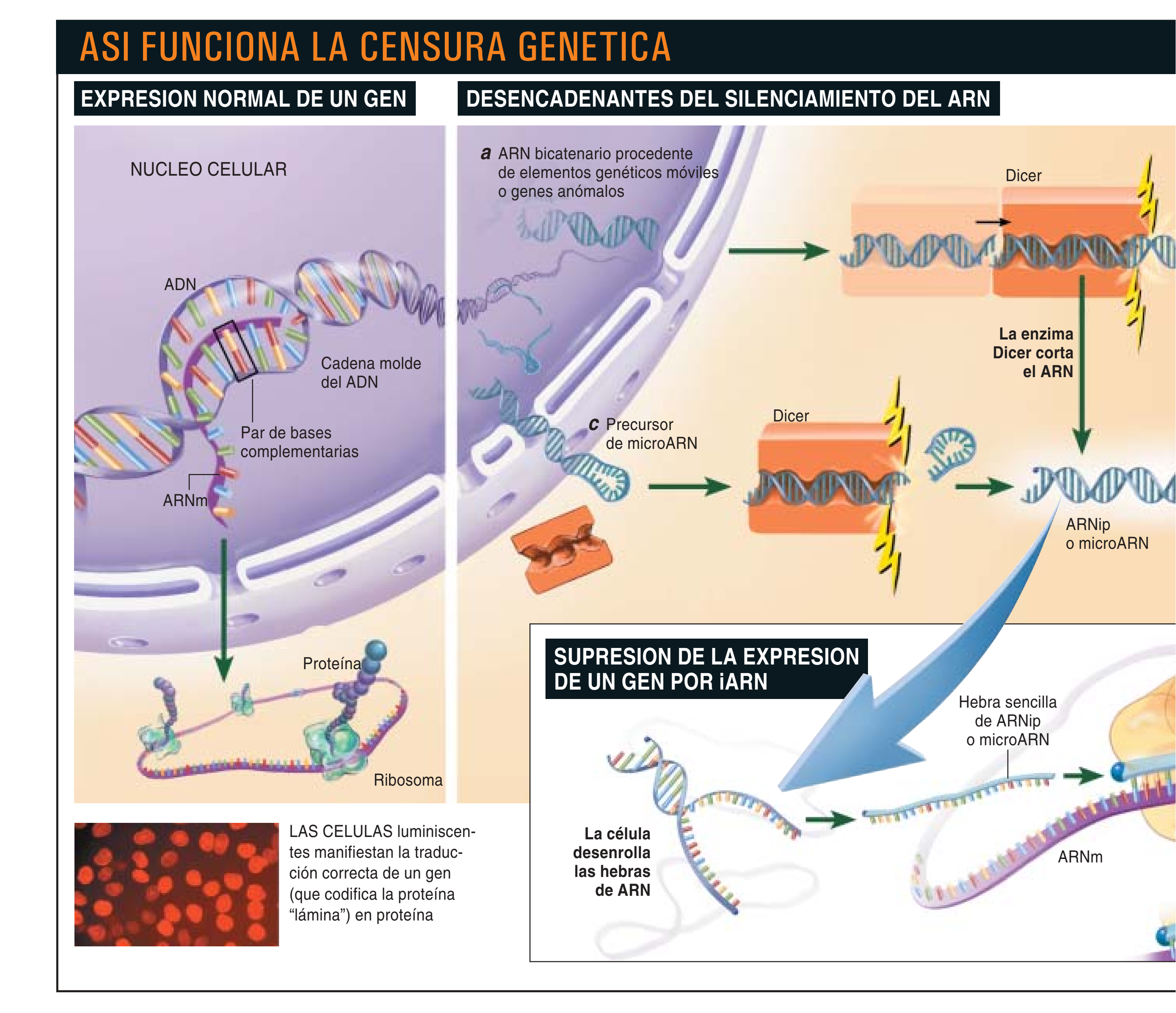


2. LOS NEMATODOS LUMINISCENTES demostraron que la interferencia de ARN operaba en los animales, no sólo en las plantas. Cuando los gusanos cuyas células expresan un gen determinante de una proteína fluorescente (*izquierda*) se tratan con ARN bicatenario correspondiente al gen, desaparece la luz (*derecha*).

CORTESIA DE LISA TIMMONS *Universidad de Kansas*; TIMMONS ET AL. EN *GENE*, VOL. 263; 2001

Los autores

NELSON C. LAU Y DAVID P. BARTEL se han ocupado del estudio de moléculas de ARN minúsculas que regulan la expresión génica. Lau trabaja en el Instituto Whitehead y en el Tecnológico de Massachusetts; en este último enseña Bartel.



LAS CELULAS luminiscentes manifiestan la traducción correcta de un gen (que codifica la proteína "lámina") en proteína

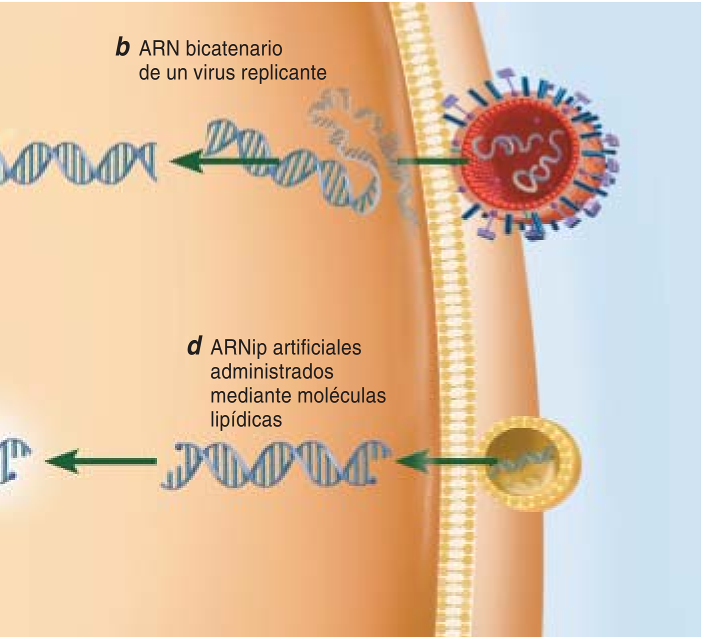
ARNip del CSIR se adherirá sólo al ARN mensajero que exactamente se complemente con su propia secuencia nucleotídica. Por tanto, a diferencia de la respuesta del interferón, el complejo silenciador es muy selectivo a la hora de escoger su ARNm objetivo.

Cuando un ARN mensajero que encaje finalmente se haya adherido al ARNip, la enzima "Slicer" ("rebanadora") parte en dos la cadena de ARN. El CSIR libera entonces los dos trozos de ARNm (incapacitados ahora para dirigir la síntesis de proteína) y prosigue en su labor. En sí mismo, el CSIR permanece intacto, libre para salir al encuentro de otro ARN mensajero y escindirlo. A través de esa vía, la iARN censora emplea trozos de ARN bicatenarios a modo de lista negra para identificar y silenciar los correspondientes ARN mensajeros.

El equipo liderado por David C. Boulcombe, del Laboratorio Sainsbury en Norwich, detectó ARNip en vegetales. Posteriormente, el grupo de Tuschl los aisló de la mosca del vinagre y demostró su papel en la silenciación de genes, tras sintetizar ARNip artificiales y usarlos para dirigir la destrucción de ARNm dianas. Cuando lo consiguió, Tuschl se planteó si estos trocitos de ARN podrían deslizarse en las células de los mamíferos sin suprimir la red de alerta de la respuesta del interferón, que habitualmente se desentiende de los ARN bicatenarios de menos de 30 pares de nucleótidos. Introdujo ARNip sintéticos en cultivos de células de mamíferos. El experimento ratificó lo esperado. Se silenciaron los genes diana; nunca se activó la respuesta de interferón.

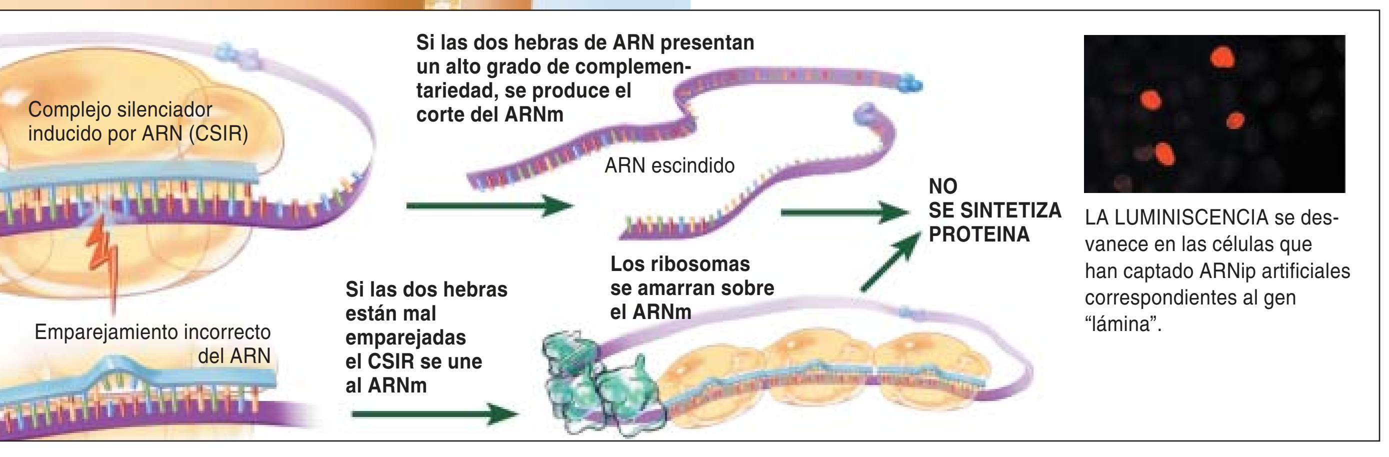
Los éxitos de Tuschl sirvieron de acicate para muchos investigadores. Desde hacía tiempo, los genéticos habían conseguido introducir genes nuevos en las células de mamíferos, sirviéndose de vectores víricos. Pero bloquear un gen de interés para determinar su

TERESE WINSLOW; THOMAS TUSCHL *Universidad Rockefeller (micrografías)*



UNA CELULA CENSURA la expresión de un determinado gen cuando se interpone en el camino del ARN mensajero (ARNm) transcrito a partir del gen dañino; evita con ello que los ribosomas traduzcan el ARN en proteína activa según lo habitual (*panel izquierdo*). De la puesta en marcha de la maquinaria de censura encargan moléculas pequeñas de ARN bicatenario que portan sueltos los extremos. De los ARN bicatenarios, largos y producidos al autocopiarse secuencias génicas (*a*) o virus (*b*), Dicer, una enzima tajadora, trocea cortos ARNip. Los precursores de microARN (*c*), secuencias de ARN reguladoras, resultan también de la acción de la mencionada enzima. Así se nos faculta para, luego, utilizando moléculas lipídicas, insertar ARNip artificiales en las células (*d*).

Los fragmentos de ARN se separan en cadenas sueltas (*panel inferior*), que se combinan con proteínas para formar un complejo silenciador inducido por ARN (CSIR). El CSIR aprehende entonces el ARNm que sea complementario de la breve secuencia de ARN. Si el emparejamiento es realmente perfecto el mensaje secuestrado se trocea en fragmentos inútiles (*hilera superior*). Pero un emparejamiento menos cabal insta una respuesta diferente; por ejemplo, que el CSIR bloquee los movimientos ribosómicos y, con ello, suspender la traducción del mensaje en proteína (*hilera inferior*).



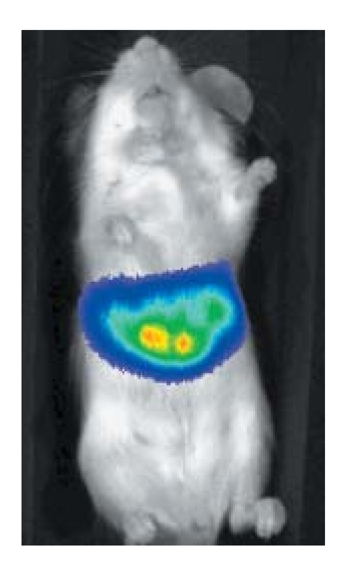
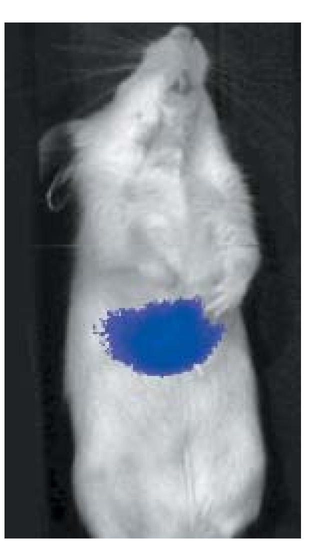
LA LUMINISCENCIA se desvanece en las células que han captado ARNip artificiales correspondientes al gen "lámina".

función propia podría llevar meses de trabajo en el laboratorio. Ahora, el sueño dorado de silenciar con facilidad un gen particular en el interior de las células de mamíferos se convertía, de la noche a la mañana, en una realidad al alcance de la mano. En cuestión de horas, con los ARNip puede desactivarse casi cualquier gen de interés en un cultivo celular, incluidas las líneas celulares humanas. Y, además, el efecto persiste durante días, lo suficiente para completar un experimento.

Una herramienta soñada

Si importante resulta la interferencia de ARN para los biólogos de mamíferos, mayor servicio rinde a quienes investigan en organismos inferiores. Los estudiosos de gusanos y vegetales se encuentran con una ventaja añadida: en estos organismos, el efecto censor se extiende y difunde más allá del punto en que se introdujo el ARN bicatenario. En razón de ese carácter sistémico, se ha aprovechado en gusanos las posibilidades que ofrece la iARN, con sólo alimentarlos de bacterias genéticamente manipuladas para que sinteticen el ARN bicatenario correspondiente al gen que se desea acallar.

Fácil de inducir y poderosa en su acción, la interferencia de ARN aviva la imaginación de los científicos. Secuenciados ya genomas enteros de diversos organismos, podemos recurrir a la interferencia de ARN para explorar, de una manera sistemática, qué sucede cuando se silencia un gen. Es lo que acaban de realizar cuatro grupos en millares de experimentos paralelos; cada uno se encargó de inactivar, en *C. elegans*, un gen diferente. Se ha emprendido un estudio similar de genomas enteros de plantas; varios equipos se

3. SI SE LES INYECTA ADN que contiene el gen de la luciferasa, los ratones devienen fotoemisores (*izquierda*). Pero ese efecto se desvanece si, por el contrario, se les inyecta ARNip correspondiente al gen (*derecha*); se evidencia así una nueva vía para extraer todo el partido posible de la iARN en mamíferos.

han aunado para acometer una investigación exhaustiva de la iARN en células de mamíferos.

El potencial que encierra la interferencia de ARN no se le ha escapado a la industria farmacéutica. En algunos laboratorios se recurre a ese fenómeno para rastrear todos los genes de una clase determinada, en busca de objetivos prometedores para nuevos medicamentos. El silenciamiento sistemático de genes mediante la iARN podría permitir descubrir un gen imprescindible para el desarrollo de ciertas células cancerosas, pero menos importante para el de las células normales. Podría entonces crearse un candidato farmacológico que se interpusiera en la síntesis de la proteína de dicho gen y contrastar así la eficacia del fármaco contra el cáncer. Los laboratorios han apostado fuerte sobre la posibilidad de que el silenciamiento de un gen por iARN se convierta en terapia contra el cáncer, las infecciones víricas, ciertas alteraciones genéticas dominantes y otras enfermedades que podrían controlarse al evitar que los genes implicados sinteticen las proteínas culpables de la patología.

En numerosas publicaciones se apunta ya en esa dirección. Al menos seis laboratorios han conseguido, en cultivos de células humanas, detener transitoriamente la proliferación del virus VIH, el de la poliomielitis y el de la hepatitis C. Se expusieron las células a ARNip; las células interrumpieron entonces la síntesis de proteínas cruciales para la reproducción de los agentes patógenos. Más recientemente, los grupos dirigidos por Judy Lieberman, de la facultad de medicina de Harvard, y Mark A. Kay, de la facultad de medicina de la Universidad de Stanford, han dado a conocer que los ARNip inyectados en ratones, en condiciones de presiones extremadamente elevadas, frenaron hepatitis y libraron a muchos de una enfermedad hepática que hubiera acabado con su vida.

A pesar de estos éxitos de laboratorio, habrán de pasar años antes de que terapias basadas en la iARN puedan aplicarse en los hospitales. El reto más difícil será posiblemente el de la administración. Aunque el efecto de iARN puede diseminarse en una planta o un gusano, tal propagación no parece que ocurra en el hombre y otros mamíferos. Añádase que los ARNip son muy voluminosos si se comparan con los fármacos al uso; no pueden tomarse como comprimidos, porque el sistema digestivo los destruiría antes de que se absorbieran. Se están ensayando diversas vías para diseminar los ARNip por distintos órganos y conducirlos a través de las membranas externas de las células.

Otro método para resolver el problema de la administración lo aporta la terapia génica. Un gen nuevo que produzca un ARNip particular podría cargarse en el interior de un virus benigno que lo llevaría luego hasta el interior de las células que infectara. A este respecto, el grupo de Beverly Davidson, de la Universidad de Iowa, se ha valido de un adenovirus modificado para introducir genes productores de ARNip en el cerebro e hígado de ratones. Pero la terapia génica aplicada al hombre se enfrenta hoy con dificultades técnicas y normativas.

Si dejamos de lado los problemas relacionados con la administración, es obvio que los nuevos planteamientos relacionados con la iARN han levantado un entusiasmo que no despertaron las técnicas del ARN catalítico, ni el ARN antisentido; ambos métodos podrían usarse, en principio, con fines terapéuticos al impedir la intervención de ARN mensajero nocivos. Esa esperanza se apoya, en parte, en el reconocimiento de que la interferencia de ARN domeña la maquinaria natural de censura de genes que la evolución ha ido perfeccionando en el transcurso del tiempo.

¿Por qué tienen censores las células?

Se cree, en efecto, que el mecanismo de la censura de genes emergió hace unos mil millones de años para proteger algunos precursores de vegetales, animales y hongos contra virus y elementos genéticos móviles. Respaldan esa tesis los trabajos de Ronald H. A. Plasterk, del Instituto Holandés del Cáncer, y de Hervé Vaucherer, del Instituto Nacional Francés de Investigaciones Agrarias. Demostraron que los gusanos modernos cuentan con la interferencia de ARN para defenderse contra los elementos genéticos móviles y, las plantas, contra los virus.

Pero la interferencia del ARN cumple también otras funciones biológicas. Los gusanos y las malas hierbas mutantes, con la enzima Dicer alterada o en cuantía insuficiente, sufren numerosos defectos de desarrollo y no pueden reproducirse. ¿Por qué acarrea alteraciones tan profundas, en animales y plantas, una merma de la enzima Dicer?

De acuerdo con cierta hipótesis, una vez que la naturaleza adquirió un mecanismo tan eficaz para el si-

ANTON P. McCAFFREY ET AL. EN *NATURE*, VOL. 418; 4 de julio, 2002; SISTEMA DE CREACION DE IMAGENES CORTESIA DE CHRISTOPHER CONTAG

lenciamiento de genes subversivos en virus y secuencias móviles de ADN, comenzó a pedir en préstamo las virtualidades de la iARN y aplicarlas a diversos fines. Cada célula tiene el mismo conjunto de genes; lo que distingue a una de otra son los genes que se han expresado y los que no. La mayoría de las plantas y de los animales arrancan de una célula embrionaria, que se divide y termina por dar lugar a una multitud de células de diversos tipos. Para que esto ocurra, muchos de los genes que se han expresado en las células embrionarias deben inactivarse conforme va madurando el órgano. Otros genes, hasta entonces inactivos, entran en funcionamiento. Cuando la maquinaria de la iARN no está ocupada en la defensa contra el ataque, interviene manifiestamente con su ayuda en el silenciamiento de genes celulares normales durante las transiciones del desarrollo, necesarias para formar tipos celulares dispares (neuronas o miocitos, por ejemplo) u órganos diferentes (cerebro y corazón).

¿Qué es entonces lo que mueve a la maquinaria de la iARN a acallar determinados genes normales en el interior de la célula? En algunos casos, una célula puede producir ARN bicatenario, largo, con ese fin específico. A menudo, sin embargo, los estímulos disparadores son “microARN” (pequeños fragmentos de ARN semejantes a los ARNip, aunque distintos por su origen). Mientras que los ARNip proceden de los mismos tipos de genes o regiones genómicas que terminan por quedar silenciados, los microARN proceden de genes cuya única misión es producir estos ARN reguladores diminutos.

La molécula de ARN inicialmente transcrita de un gen de microARN —el precursor del microARN— se pliega sobre sí misma, formando una estructura similar a una horquilla. Con la intervención de la enzima Dicer, se rebana de la horquilla la sección media; la pieza resultante se comporta a la manera de un ARNip, con una salvedad importante: no censurará ningún gen que semeje al que lo ha producido, pero sí a otro distinto.

Igual que sucedió con el fenómeno de la iARN en general, los biólogos tardaron en darse cuenta del potencial que encerraban los microARN para regular la expresión génica. Hasta hace poco, sólo tenían noticia de dos microARN, el *lin-4* ARN y el *let-7* ARN, descubiertos por los grupos de Victor Ambros, de la facultad de medicina de Dartmouth, y Gary Ruvkun, de la facultad de medicina de Harvard. En los dos últimos años, nosotros, Tuschl, Ambros y otros hemos descubierto centenares de genes adicionales de microARN en gusanos, plantas y hombre.

Con Christopher Burge, del MIT, hemos estimado que el hombre tiene de 200 a 255 genes de microARN; ello significa casi un 1 por ciento del número total de nuestros genes. Los genes de microARN se habían escapado a la detección porque los programas informáticos diseñados para examinar montones de datos de secuencias genómicas no se habían afinado para este tipo insólito de gen, cuyo producto final es un ARN, no una proteína.

Algunos microARN, en particular los de las plantas, dirigen el proceso de corte y empalme de sus ARNm dianas. Lo demostraron James C. Carrington, de la Universidad estatal de Oregón, y Zamore. Nosotros y Bonnie Bartel, de la Universidad Rice, hemos comprobado que los microARN vegetales se ordenan fundamentalmente hacia genes que intervienen en el desarrollo. Al limpiar sus mensajes procedentes de ciertas células durante el desarrollo, las iARN facilitan que las células maduren en el tipo correcto y formen las estructuras pertinentes.

Merece subrayarse que los ARN *lin-4* y *let-7*, cuya existencia se descubrió en los gusanos en razón de su papel decisivo para pautar el desarrollo, pueden emplear también una segunda táctica. Los ARNm que constituyen la diana de estos microARN, son sólo aproximadamente complementarios de los microARN; tales mensajes no son escindidos. Existe otro mecanismo que bloquea la traducción de los ARN mensajero en proteínas funcionales.

A la vista de estos mecanismos diferentes de silenciamiento, los biólogos mantienen fija la atención sobre las funciones de los ARN pequeños y de la maquinaria de la iARN. Se multiplican los datos indicadores de que los ARNip no sólo atrapan ARN mensajeros para su destrucción, sino que además dirigen el silenciamiento de ADN (en el caso más extremo, eliminando los genes del genoma). En la mayoría de los casos, sin embargo, el ADN silenciado no se destruye; sencillamente, experimenta un empaquetamiento más denso para que no pueda transcribirse.

Desde unos comienzos humildes en flores blancas y gusanos contrahechos, nuestras ideas acerca de la interferencia de ARN han recorrido un largo trecho. Casi todas las facetas de la biología, biomedicina y bioingeniería comienzan a verse afectadas por las iARN, al par que la técnica del silenciamiento de genes se extiende por nuevos laboratorios y organismos experimentales.

Más aún, la iARN nos plantea muchas cuestiones fascinantes. ¿En qué procesos biológicos repercuten la interferencia de ARN, los ARNip y los microARN? ¿De qué modo los mecanismos moleculares de la iARN operan en el plano atómico y en el de enlace químico? ¿Qué enfermedades cabe atribuir a deficiencias en la iARN o a los procesos de microARN? A medida que se vayan despejando estos interrogantes, nuestras ideas sobre el fenómeno se irán asentando hasta constituir un sólido fundamento para la medicina genética.

Bibliografía complementaria

RNAI: NATURE ABHORS A DOUBLE STRAND. György Hutvágner y Philip D. Zamore en *Current Opinion in Genetics & Development*, vol. 12, n.º 2, págs. 225-232; abril de 2002.

GENE SILENCING IN MAMMALS BY SMALL INTERFERING RNA*s*. Michael T. McManus y Philip A. Sharp en *Nature Reviews Genetics*, vol. 3, págs. 737-747; octubre de 2002.

MICRORNA*s*: AT THE ROOT OF PLANT DEVELOPMENT? Bonnie Bartel y David P. Bartel en *Plant Physiology*, vol. 132, n.º 2; págs. 709-717; junio de 2003.

Informatización y desarrollo humano

En la difusión de la informática no cuenta tan sólo la mera disyuntiva entre tener y carecer

Mark Warschauer

Desde hace unos años ha ido preocupando cada vez más, a los gestores públicos y a los sociólogos, una posible escisión de la sociedad entre quienes disponen y quienes carecen de acceso a los ordenadores y a Internet. La Administración Nacional de la Información y las Telecomunicaciones de EE.UU. acuñó a mediados del decenio de 1990 para referirse a esta situación una etiqueta que se hizo popular: la "divisoria digital". No se tardó en describir con ella la difusión de la informática en unos y otros países.

Las disparidades subyacentes son reales, tanto en el seno de cada país como entre naciones distintas. La Fundación Benton, que promociona el uso público de las telecomunicaciones, informa de que, a finales del año 2001, el 80 por ciento de las familias estadounidenses con una renta total mayor de 75.000 dólares tenía una conexión a la Red, frente a sólo un 25 por ciento de las familias más pobres. El 55 por ciento de los blancos, el 31 por ciento de los afroamericanos y el 32 por ciento de los hispanos accedían a Internet desde sus domicilios. En la mayor parte de los países africanos, menos del 1 por ciento de la población está conectado a Internet. Como es de imaginar, tal disparidad informática se corresponde con otras desigualdades sociales y económicas.

Ahora bien, la palabra "divisoria", binaria y simplista, no hace justicia a las muy variadas formas en que personas y pueblos acceden a Internet y la usan. Quien navegase por la Red una vez al mes desde la biblioteca pública de una localidad norteamericana caería del lado de los "desprovistos", mientras que en un país subdesarrollado se contaría entre "los que tienen". Ver la cuestión sólo en blancos y negros puros puede llevar a quienes intenten paliar las desigualdades tecnológicas por un camino equivocado. Rob Kling, que fue director del Centro de Informática Social de la Universidad de Indiana, lo enunció con claridad: "[El] gran problema del planteamiento implícito en la expresión 'divisoria digital' es que sobreentiende que también las soluciones son 'digitales', es decir, que consisten en distribuir ordenadores y sistemas de telecomunicación sin tomar en consideración los contextos en que se vaya a instalarlos".

TOM DRAPER DESIGN



DE IZQUIERDA A DERECHA: INSTITUTO NACIONAL DE TECNOLOGÍA DE LA INFORMACION, DAVID GREEDY, PHOTODISC, PHOTODISC (*línea superior*); PHOTODISC (*segunda línea*); PHOTODISC, PHOTODISC, MARK WARSCHAUER (*tercera línea*); PHOTODISC, PHOTODISC, KEITH DANNEMILLER *Corbis/SABA* (*cuarta línea*); KEITH DANNEMILLER *Corbis/SABA*, INSTITUTO NACIONAL DE TECNOLOGÍA DE LA INFORMACION, PHOTODISC, PHOTODISC (*quinta línea*); PHOTODISC, PHOTODISC, VIVIAN MOOS *Corbis* (*línea inferior*)

No han faltado quienes hayan supuesto, conforme a esa creencia, que la penuria de acceso digital que padecen naciones, comunidades e individuos podría subsanarse fácilmente con un reparto generoso de ordenadores y conexiones a Internet. Newt Gingrich, ex presidente de la Cámara de Representantes de EE.UU., ha hablado de los beneficios que se derivarían de proporcionar un ordenador portátil a todos los niños de su país, pero sin ofrecer al mismo tiempo un plan bien concebido para su utilización. Bill Gates ha donado ordenadores a las bibliotecas públicas de pueblecitos por todo Estados Unidos, convencido de que la conexión a Internet contribuirá a frenar el éxodo de población de las zonas rurales a las urbanas. Cierto es que ese servicio ha mejorado la vida de los beneficiados al facilitarles el contacto con amigos y parientes; pero no ha frenado el éxodo, que depende de factores de mayor empaque, como la existencia de puestos de trabajo. Puede incluso que haya contribuido a él, al facilitar la búsqueda a distancia de empleo en las ciudades. Se ha de decir, en honor de Gates y Gingrich, que al menos tuvieron presente que el acceso a la técnica es un problema. Gates, consciente de las limitaciones de la informática para resolver los males sociales, ha donado desde entonces miles de millones de dólares a campañas más amplias, de carácter educativo o sanitario, en todo el mundo.

El punto de vista implícito en esas actitudes se denomina, en ciertos círculos académicos, determinismo tecnológico: la creencia en que la mera presencia de la técnica conduce a sus aplicaciones ordinarias, las cuales, a su vez, traen consigo el cambio social. Christopher Dede, de la Escuela de Posgraduados en Educación de la Universidad de Harvard, ha dado en hablar al respecto del "modelo fuego", según el cual el ordenador, con su mera presencia, engendra saber o desarrollo, tal y como el fuego calienta a su alrededor. Y así, los gobiernos, el sector privado y las fundaciones e instituciones de beneficencia han gastado cientos de millones de euros en tender puentes que salven la "divisoria digital" construidos con ordenadores y líneas de Internet, que se ofrecen a los necesitados a menudo sin prestar atención suficiente a los contextos sociales en que tendrían que utilizarse. (Dede señala que un símil mejor que el del fuego sería el de la ropa, que también nos mantiene calientes, pero está adaptada al uso y talla de cada uno.)

¿Qué resultados ha tenido en la práctica esta aplicación del determinismo tecnológico? A lo largo de estos últimos años he estado viajando por todo el mundo con el fin de estudiar programas tecnológicos de carácter comunitario, tanto en países desarrollados como en países en desarrollo. He examinado multitud de planes diversos; me he entrevistado con centenares de organizadores y de participantes. Como irán haciendo ver los casos siguientes, han quedado de manifiesto dos hechos básicos: que planes bienintencionados se desvían a menudo en direcciones inesperadas y que los fracasos más serios se producen al tratar de abordar problemas sociales complejos con un enfoque miope, basado en la mera provisión de equipos.

Resumen/*Lógica tecnológica*

- La noción de que existe una "divisoria digital" entre quienes disponen y quienes carecen de ordenadores y acceso a las telecomunicaciones es simplista y conduce a soluciones bienintencionadas, aunque incompletas, como la mera introducción de medios técnicos en unas circunstancias dadas.
- Son muy variadas las oportunidades de acceder a la informática y las telecomunicaciones y las razones para querer un nivel u otro de acceso.
- El análisis de cómo cabe servirse de los ordenadores e Internet para promover el desarrollo social reviste máxima importancia en cualquier tentativa de implantación de nuevas técnicas en un ambiente que carezca de ellas.

El autor

MARK WARSCHAUER es vicedirector del departamento de educación de la Universidad de California en Irvine, fundador y director de la revista *Language Learning & Technology*, y autor o recopilador de siete libros sobre educación, tecnología y desarrollo. Ha estudiado la repercusión de la informática en el desarrollo humano en Egipto, China, India, Brasil, Singapur y otros países.

Una metodología minimalista

En el año 2000, el gobierno municipal de Nueva Delhi, en colaboración con el Instituto Nacional de Tecnología de la Información indio, puso en marcha un experimento concebido para que los niños de uno de los barrios más pobres pudieran usar ordenadores. Los funcionarios municipales y los representantes del Instituto montaron en la calle un punto informático con cinco puestos. Los ordenadores, que se conectaban a Internet por vía telefónica, se instalaron en el interior de una cabina cerrada, pero las pantallas y los mandos sobresalían por unos huecos abiertos en una pared. Conforme a la llamada "educación mínimamente invasiva", no había ni maestros ni instructores. Se trataba de permitir a los niños el acceso diario y sin cortapisas, de modo que pudieran aprender a su propio ritmo, en lugar de hacerlo como se lo dictasen los adultos.

Los organizadores proclamaron que este programa abría nuevas vías para llevar las técnicas de la información a los pobres de las ciudades de todo el mundo. Circularon por Internet historias alentadoras acerca de niños analfabetos que aprendían por sí solos a manejar ordenadores, derribando así las barreras que les cerraban el paso hacia la era de la información.

Pero el cuadro que me hice al visitar el lugar fue muy diferente.

La conexión a Internet raramente funcionaba. Que los ordenadores se utilizasen desde una pared, como un cajero automático, en vez de en una sala, dificultaba la comunicación o la instrucción. La mayoría de los vecindarios pobres de Nueva Delhi cuentan ya con organizaciones comunitarias que trabajan con niños; podrían haberlos formado en un tipo distinto de centro informático, pero su participación no fue ni solicitada, ni bien recibida. Es cierto que a lo largo de los nueve meses que duró el experimento los niños aprendieron a manipular mandos (*joysticks)* y botones. Pero careciendo de programas educativos, y estando los contenidos en inglés y no en hindú, lo que con mayor frecuencia hicieron fue lo que habría sido de esperar: entretenerse con diversos pasatiempos y servirse de programas gráficos para dibujar.

Entre los padres hubo división de opiniones. Algunos aplaudieron la iniciativa, pero la mayoría expresó su preocupación por el hecho de que no hubiera una instrucción organizada. Incluso hubo quien se quejó de que el ordenador era contraproducente. “Mi hijo iba muy bien en la escuela”, explicaba un padre, “pero ahora se pasa todo el tiempo libre jugando en el ordenador del punto informático y sus deberes escolares se resienten.” En breve: la comunidad vino a comprender que la educación mínimamente invasiva era, en la práctica, una educación mínimamente eficaz.

A pesar de ello, los proyectos de educación técnica de todo el mundo a menudo insisten demasiado en los elementos materiales de la informática y, en cambio, se olvidan casi del marco pedagógico y curricular que conforma su uso. En el mundo académico ha salido al paso de este determinismo tecnológico la “informática social”; según esta concepción, hay que enmarcar la técnica en un contexto específico que tenga en cuenta los equipos, los programas, los recursos de apoyo y la infraestructura, sin olvidar las personas en sus diversos roles y relaciones, tanto entre sí como con otros elementos del sistema. Y la técnica y el sistema social se amoldan una al otro sin cesar, mutua y recíprocamente, lo mismo que una comunidad biológica y su entorno.



PABLO BARTHOLOMEW *Netphotograph.com*

1. ESTOS NIÑOS DE NUEVA DELHI experimentaban en el año 2000 con ordenadores que manejaban a través de aberturas en una pared. Aquel proyecto de educación mínimamente invasiva se proponía la introducción de la informática en el entorno cotidiano de los niños, de modo que pudieran aprender sin guía a valerse del ordenador. Pero carentes de instructor, el ordenador fue para ellos poco más que un juguete técnicamente avanzado.

Aunque los maestros tradicionales, los padres o los asistentes sociales puedan no estar familiarizados con el término académico “informática social”, son muchos quienes aprecian ya las consecuencias de la interrelación de la técnica y las organizaciones públicas. La informática social ha dado nacimiento recientemente a la “informática comunitaria”, que considera también aspectos peculiares de la cultura concreta en que se ha difundido la técnica, de manera que la comunidad pueda valerse con máxima eficacia de ella para alcanzar objetivos políticos, económicos o culturales.

Una experiencia más congruente

Entre los ejemplos de programas inspirados por la noción de informática comunitaria se cuenta el proyecto Gyandoot (que se traduciría como “proveedor de conocimiento”). En el año 2000, en el extremo sudoeste de Madhya Pradesh, uno de los estados más pobres de la India, el gobierno puso en marcha ese ensayo informático, que perseguía proporcionar mayor poder económico y político a la población rural, constituida casi en sus dos terceras partes por personas analfabetas y subalimentadas. Se instaló en cada aldea un centro informático, conectado en red con los demás. Se encargan de atender los aparatos empresarios de la zona. Un pequeño equipo profesional, contratado por el gobierno, crea contenidos para la intranet Gyandoot (la red que conecta los ordenadores del proyecto); para ello, se basan en un análisis de las necesidades sociales y económicas de los destinatarios de esta experiencia.

Entre los contenidos figuran los precios actualizados de los cultivos habituales en el distrito, tanto en los mercados regionales como nacionales, de modo que los pequeños agricultores puedan decidir si les conviene recoger su cosecha, y dónde venderla, sin tener que perder un día viajando hasta la capital del distrito para conocer los precios. Un servicio de reclamaciones permite a los aldeanos informar de problemas locales, como el deficiente funcionamiento de las bombas manuales o el absentismo de los maestros

de escuela. Gracias a esta nueva capacidad de los aldeanos de expresar rápidamente sus preocupaciones por medios digitales, los servicios gubernamentales han empezado a mejorar.

Manejan los ordenadores los encargados de los centros informáticos, lo que permite que el servicio, que cuesta sólo unos céntimos, sea útil incluso para los analfabetos. Los encargados ofrecen también instrucción informática a los niños de la aldea por una pequeña cuota; se aportan así ingresos adicionales a los encargados a la vez que mejora la destreza informática colectiva. Gyandoot conecta también a sus usuarios con otras iniciativas socioeconómicas más generales, como la campaña del "niño más sano", que informa sobre vacunas y nutrición.

El lanzamiento de Gyandoot ha sido poco oneroso, pues sólo utiliza un ordenador por aldea y en parte se autofinancia, ya que los operadores de los centros recuperan parte de sus costos cobrando a los usuarios unas cantidades modestas. La manera de proceder —anteponer la atención de las necesidades de los usuarios, y satisfacerlas por medio de proveedores de servicios a pequeña escala, gestionados localmente— tiene mucho en común con un sistema anterior de locutorios telefónicos que contribuyó a extender el servicio telefónico a gran parte de la India. En los nueve meses contados a partir de octubre de 2001, los centros Gyandoot han recibido unos 21.300 usuarios, el 80 por ciento de los cuales cuentan con ingresos anuales inferiores a 300 euros. Se trata de un pequeño porcentaje de la población, pero los beneficios del proyecto —así la mejora de los servicios gubernamentales— acaban por propagarse a los amigos, las familias y los compañeros de trabajo.

Está por determinar la magnitud del éxito de Gyandoot. Pero su metodología —la combinación de una distribución bien planificada de técnica, por una parte, y de una creación de contenidos y de campañas orientadas al desarrollo social por la otra— constituye sin duda una saludable alternativa a los proyectos que se basan en sembrar ordenadores y esperar que crezca algo.

Afinación de un programa en California

Otra iniciativa, ésta norteamericana, que investigué con la doctoranda Jodie Wales pone igualmente de manifiesto la importancia de la informática comunitaria. Los institutos de educación secundaria de California ofrecen a sus alumnos unos cursos (llamados de "Advanced Placement", AP) que conceden créditos de titulaciones superiores y pueden facilitarles el ingreso en las mejores universidades. Sin embargo, estos cursos están disponibles en números clamorosamente desiguales, que dependen, sobre todo, de la composición socioeconómica y étnica de las poblaciones estudiantiles. Por ejemplo, el instituto de Beverly Hills, donde los hispanos y los negros suman un 9 por ciento del total, ofreció 45 cursos de ese tipo en 1999, mientras que el de Inglewood, situado en otra parte de la misma área metropolitana, con un 97 por ciento de estudiantes negros o hispanos, ofrecía sólo tres.

Para abordar este problema, la Universidad de California inició una colaboración con los institutos del distrito de Anaheim, que cuenta con una numerosa población de origen latino. La primera tentativa consistió en un curso AP de macroeconomía "en línea", pues muchos de los estudiantes del distrito, incluso los pobres, disponían de cierto acceso a ordenadores y a Internet fuera de la escuela. Se inscribieron alumnos de varios institutos, con lo que, al menos en teoría, se había vencido el obstáculo que suponen los grupos pequeños y dispersos de estudiantes avanzados. El resultado: sólo 6 de 22 estudiantes completaron el curso. Las entrevistas y encuestas con los alumnos dejaron claras algunas de las razones: el formato de la instrucción en línea —en la que el estudiante ha de realizar las tareas en su casa, con su propio ordenador— estaba poco estructurado, y el contacto con los profesores y la interacción entre compañeros resultaban demasiado escasos como para mantener la motivación de los estudiantes, que habían de vérselas con un material difícil. Los latinos, sobre todo, comentaron que preferían disfrutar de esos apoyos sociales.

Aun así, el fallo fue fructífero. Al año siguiente, un plan revisado reunió a alumnos de varios institutos en un laboratorio de informática, esta vez con una localización central, para seguir un curso optativo de introducción a la informática y al lenguaje C de programación. Aunque se seguía impartiendo en línea, con un instructor que no se encontraba presente allí y un plan de estudios basado en el trabajo con ordenador, en el aula había un tutor que respondía preguntas y prestaba ayuda. La combinación de la enseñanza a distancia y la interacción personal con el tutor y los compañeros demostró ser mucho más eficaz: 56 de los 65 alumnos completaron el curso. Basándose en estos datos, el plan de preparación para los estudios superiores de la Universidad de California abandonó su anterior modelo de instrucción puramente "en línea", trocándolo por el modelo mixto de enseñanza a distancia y contacto personal. (Claro está, los alumnos quizás encontrasen más sencilla una clase sobre ordenadores que un curso de macroeconomía, o tal vez resulte más idónea para la enseñanza en línea. Es necesario considerar tales aspectos cuando se establezca la modalidad de un curso.)

Cambiar la mentalidad

Son cada vez más las pruebas que señalan la necesidad de una consideración cuidadosa de todas las posibles ramificaciones antes de aplicar la técnica como si se tratase de un ungüento amarillo educativo. Mis investigaciones, junto con las de otros pedagogos como Henry Becker, de la Universidad de California en Irvine, Harold Wenglinsky, de la Universidad de la Ciudad de Nueva York, y Janet Schofield, de la Universidad de Pittsburgh, indican que la introducción de ordenadores en las escuelas tiene tantas posibilidades de exacerbar las desigualdades como de mitigarlas. La clave no estriba

en el desigual acceso a la informática, sino en la muy diferente forma en que es utilizada. Nuestros trabajos señalan que los alumnos de nivel socioeconómico alto se valen con mayor frecuencia, desde preescolar hasta la entrada en la secundaria, de los ordenadores para experimentar e investigar, para la indagación crítica, mientras que los estudiantes pobres participan en tareas repetitivas y ejercicios menos difíciles, en los que no se saca partido pleno de la informática. En las clases de matemáticas y de lengua, donde son corrientes este tipo de pruebas reiterativas, los estudiantes pobres utilizan más los ordenadores que los más pudientes. Tan sólo en las clases de ciencias que se basen en experimentos y simulaciones utilizan los estudiantes ricos más a menudo los ordenadores que los pobres. Una vez más, un planteamiento de la "divisoria digital" que tan sólo tenga en cuenta la mera facilidad de acceso dejará a un lado estas desigualdades más generales en el uso y el aprendizaje de la técnica.

Las personas accedemos a la información digital de muy diversos modos, por lo común desde el interior de redes sociales tendidas entre parientes, amigos y compañeros de trabajo. El aprendizaje de la lectura nos proporciona una analogía adecuada. No existe una "divisoria bipolar" que ponga a un lado a quienes saben leer lo que sea y al otro a quienes no saben hacerlo en absoluto. Antes bien, hay una serie de niveles de lectura, relacionados con los distintos propósitos escolares, literarios, cívicos, vocacionales y funcionales. Y la capacidad de lectura no se adquiere por mero contacto físico con los libros, sino gracias a la enseñanza, la comunicación, las relaciones laborales, el apoyo familiar y la ayuda de las redes sociales. De forma similar, cabe aplicar la informática para aumentar y mejorar los programas e iniciativas sociales ya existentes.

El saldo es, en definitiva, que no existe una divisoria digital binaria, como no existe un factor único que la determine, o permita salvarla. La técnica no es una variable exógena que pueda introducirse desde el exterior para inducir ciertos resultados. Va interrelacionada con los sistemas y procesos sociales. Y en cuanto a la política a seguir, el objetivo de llevar la técnica a los grupos marginados no consiste meramente en superar una divisoria informática, sino en fomentar la inclusión social. Para alcanzar esa meta no basta con proporcionar ordenadores y enlaces a Internet, o valerse de plataformas en línea. Debemos también presentar contenidos adecuados en diversas lenguas, promover la alfabetización y la educación, y movilizar apoyos comunitarios e institucionales orientados al logro de objetivos sociales importantes. La técnica se convierte entonces en un medio, con frecuencia muy poderoso, en lugar de constituir un fin en sí misma.

Debe señalarse que la administración Bush está recortando la financiación de los proyectos que alientan el acceso a la informática. Se podría aducir que, si no existe la divisoria digital, esos recortes son oportunos, pero semejante razonamiento es tan especioso como las soluciones simplistas basadas en la idea de la divisoria. La operación inversa de la división es la *multiplicación*. Quienes confeccionan políticas deberían dejar de pensar en divisorias que salvar. La combinación de un reparto bien planificado de medios informáticos, dotados de contenidos pertinentes, con mejoras educativas y un apoyo más resuelto, puede multiplicar los recursos que los grupos sociales poseen ya.



PABLO BARTHOLOMEW *Netphotograph.com*

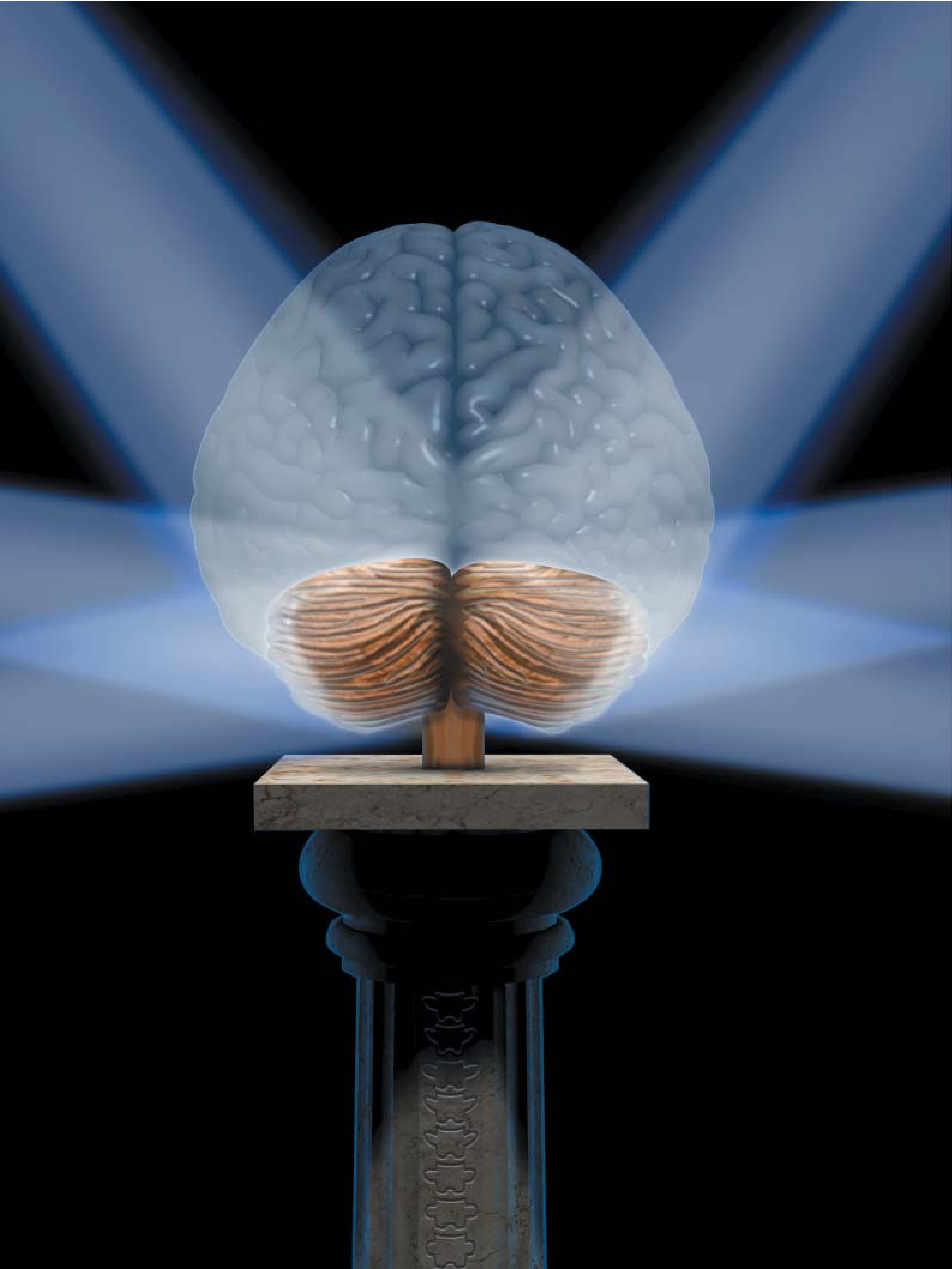
2. UNOS CAMPESINOS ACCEDEN a la intranet Gyandoot desde un centro comunitario informatizado del distrito Dhar, en India central, donde el 60 por ciento de sus 1.700.000 habitantes tienen rentas inferiores al nivel de la pobreza. Esa red les facilita los precios de las cosechas, formularios oficiales y un lugar donde celebrar subastas de aldea y desde el que expresar sus quejas.

Bibliografía complementaria

WHO'S WIRED AND WHO'S NOT? Henry J. Becker en *The Future of Children*, vol. 10, n.º 2; 2000.

RECONCEPTUALIZING THE DIGITAL DIVIDE. Mark Warschauer en *First Monday*, vol. 7; n.º 7; 2002.

TECHNOLOGY AND SOCIAL INCLUSION: RETHINKING THE DIGITAL DIVIDE. Mark Warschauer. MIT Press, 2003.



Reconsideración del cerebelo

Sabemos ahora que el cerebelo, largamente tenido por mero coordinador cerebral de los movimientos del cuerpo, se halla activo durante un extenso repertorio de actividades perceptivas y cognitivas

James M. Bower y Lawrence M. Parsons

"En la parte posterior del cráneo, encaramado sobre el tallo cerebral, envuelto y recubierto por el manto de los grandes hemisferios cerebrales, se encuentra una masa de tejido cerebral gris y blanco, en forma de habichuela y del tamaño de una pelota de frontón. Se trata del cerebelo, 'el cerebro menor'".

Así empezaba, con cierta modestia, el artículo que en 1958 presentaba el cerebelo a los lectores de *Scientific American*. Era su autor Ray S. Snider, de la Universidad Noroccidental, y la introducción seguía: "En contraste con el cerebro, donde los humanos han buscado y hallado los centros de tantas y tan esenciales actividades mentales, el cerebelo continúa siendo una región de sutil y tentador misterio, que oculta su función a los investigadores". Pero cuando apareció el segundo artículo sobre el cerebelo en *Scientific American*, 17 años más tarde, su autor, Rodolfo R. Llinás (en la actualidad, en el Centro Médico de la Universidad de Nueva York), afirmaba confiadamente: "Ya no hay duda de que el cerebelo es un enclave central de control de la organización del movimiento".

Sin embargo, la función del cerebelo ha vuelto recientemente a ser materia de debate. En particular, las ciencias de la neurocognición, valiéndose de potentes medios de formación de imágenes cerebrales, han descubierto que el cerebelo humano se halla activo durante una amplia variedad de actividades que no guardan relación directa con el movimiento. En refinados estudios de cognición se ha revelado, además, que la lesión de determinadas áreas del cerebelo puede provocar un imprevisto deterioro en procesos no motores, sobre todo, en lo tocante a la rapidez y la precisión con la que es percibida la información sensorial.

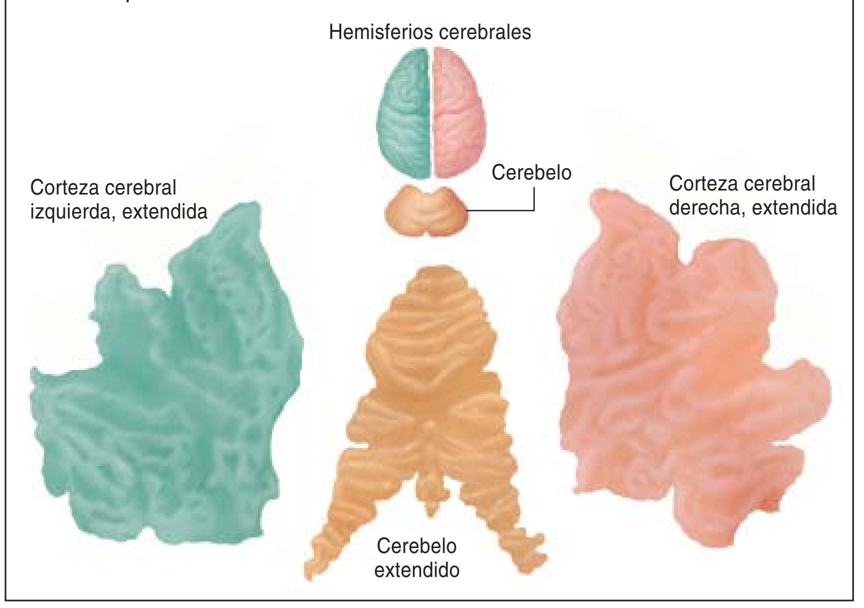
Otros hallazgos indican que el cerebelo puede desempeñar roles importantes en la memoria a corto plazo, en la atención, el control de los impulsos, las emociones, la cognición superior, la capacidad para planear y realizar tareas; participaría incluso en ciertas enfermedades y afecciones (esquizofrenia y autismo). De lo recabado en otros experimentos neurobiológicos adicionales —referidos tanto a las pautas de señales sensorias aferentes al cerebelo como a las vías por las que el cerebelo procesa tales informaciones— se sugiere, asimismo, la necesidad de revisar sustancialmente las ideas vigentes sobre la función de este órgano. El cerebelo ha vuelto a convertirse en un área de "tentador misterio".

En retrospectiva, tal vez no resulte sorprendente que el cerebelo actúe como algo más que mero coordinador del movimiento. Su volumen considerable y su intrincada estructura dan a entender que su papel es más complejo y universal. Su tamaño sólo cede ante la corteza cerebral —rugosa superficie de los dos grandes hemisferios cerebrales— donde se asientan muchas funciones cerebrales cruciales. El cerebelo, al igual que la corteza cerebral, condensa en un espacio muy reducido una prodigiosa cantidad de circuitería, replegándose sobre sí mismo numerosas veces. De hecho, el cerebelo humano está mucho más replegado que la corteza; en

FRANK IPPOLITO

MAYOR DE LO QUE PARECE

AL EXTENDER la capa externa —la corteza— de los hemisferios cerebrales, y hacer lo mismo con el cerebelo, se aprecia que éste tiene, en el hombre, la superficie de un hemisferio; cuando está replegado ocupa, sin embargo, un volumen mucho menor. El tamaño y la complejidad del cerebelo indican que su función ha de ser crucial.



diversos mamíferos, es la única estructura cerebral plegada. Al extender sobre un plano el cerebelo humano resulta una lámina cuya superficie media es de 1128 cm^2, algo mayor que la funda de un disco de vinilo. Es, desde luego, bastante más de la mitad de los 1900 cm^2 que suman las superficies de las dos cortezas cerebrales.

Es obvio que el cerebelo tiene a su cargo una tarea importante, porque ha persistido —y se ha hecho más grande— en el decurso de la evolución. Aunque los biólogos suelen considerar que el elemento característico de la evolución del cerebro humano ha sido el crecimiento de la corteza, lo cierto es que también el cerebelo se ha agrandado notoriamente. Según testimonian los cráneos fósiles, su tamaño se ha triplicado, cuando menos, durante el último millón de años de historia humana. Tal vez, sin embargo, el rasgo más característico del cerebelo estribe en el número de células nerviosas; contiene más neuronas que todo el resto del cerebro. Además, la forma en que tales neuronas se hallan interconectadas ha permanecido esencialmente constante durante más de 400 millones de años de evolución de los vertebrados. Las neuronas del cerebelo de un tiburón están organizadas en circuitos casi idénticos a los de un ser humano.

No sólo el movimiento

La hipótesis de que el cerebelo controla el movimiento fue propuesta por fisiólogos de mediados del siglo XIX, tras observar que la remoción del órgano provocaba dificultades inmediatas para la coordinación del movimiento. Durante la primera guerra mundial, Gordon Holmes aportó mucho detalle a estos primeros hallazgos, visitando las tiendas de los frentes de batalla y documentando la falta de coordinación motora en soldados que habían sufrido heridas cerebelares por bala o metralla.

Sin embargo, el refinamiento de las técnicas de ensayo aplicadas a lo largo de los últimos 15 años ha complicado el panorama. En 1989, Richard K. Ivry y Steven W. Keele, de la Universidad de Oregón, observaron que los pacientes con lesiones cerebelares se mostraban incapaces de juzgar con precisión cuánto tiempo mediaba entre dos señales sonoras. A principios del decenio subsiguiente Julie A. Fiez y su grupo, de la Universidad de Washington, comprobaron que por culpa del daño en el cerebelo, los pacientes evidenciaban mayor propensión a incurrir en errores al realizar ciertas tareas verbales. Uno de estos individuos, por ejemplo, necesitaba mucho más tiempo para pensar en un verbo adecuado, como "afeitarse", al serle ofrecido un dibujo de una navaja barbera. En cambio, tal sujeto lograba antes un descriptor como "afilada".

Los autores, en estudios más recientes, han demostrado que los pacientes que sufren enfermedades neurovegetativas que provocan específicamente la contracción del cerebelo son, a menudo, menos precisos que otros al juzgar diferencias finas en la altura de dos tonos sonoros. En la misma línea, el equipo de Peter Thier, de la Universidad de Tubinga, ha hallado que los individuos que sufrían lesiones, o contracción, en todo o en parte del cerebelo, propendían a cometer errores en pruebas donde se les solicitaba que detectaran la presencia, la ve-

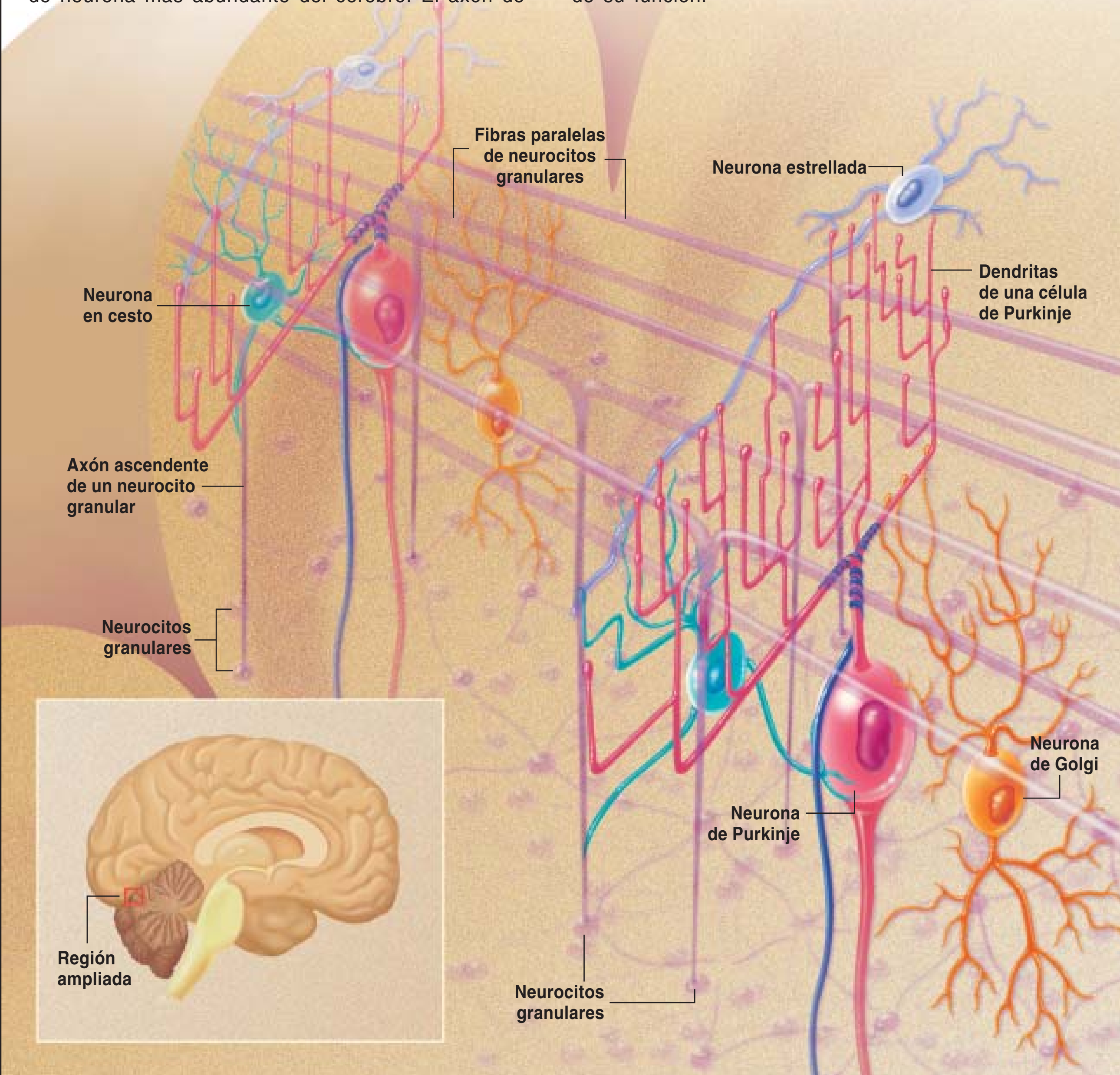
Resumen/*El cerebelo*

- El cerebelo se asienta en la base del cerebro. Posee una compleja circuitería neuronal que ha permanecido casi invariable a lo largo de la evolución de los vertebrados.
- Se cuestiona la idea tradicional de que el cerebelo se ocupa del control del movimiento. De acuerdo con la investigación reciente, se muestra activo durante un variado repertorio de tareas. El cerebelo podría incluso hallarse más implicado en la coordinación de las señales aferentes de los sentidos que en la producción de salida motora.
- A menudo, la extirpación del cerebelo en individuos jóvenes provoca escasas manifestaciones de dificultades conductuales. Ello sugiere que el cerebro puede aprender a funcionar sin cerebelo.

ALICE Y. CHEN; FUENTE: DAVID C. VAN ESSEN *Washington University School of Medicine*

ASI ESTA CONECTADO EL CEREBELO

LOS RASGOS FUNDAMENTALES de la circuitería cerebelar son conocidos desde los trabajos seminales de Santiago Ramón y Cajal, a finales del siglo XIX. La neurona central es la célula de Purkinje, así llamada en honor de Johannes E. Purkinje, que la identificó en 1837. La de Purkinje es una de las mayores neuronas del sistema nervioso y proporciona la única señal eferente de la corteza cerebelar. Recibe a través de sus 150.000 a 200.000 sinapsis un número extraordinario de señales aferentes, unas diez veces más que cualquier otra neurona de la corteza cerebral. Estas señales que les llegan proceden, sobre todo, de una de las neuronas más pequeñas de los vertebrados, la neurona granular cerebelar. Los neurocitos granulares alcanzan una astronómica densidad de empaquetamiento: seis millones por milímetro cuadrado. En razón de ello nos hallamos ante el tipo de neurona más abundante del cerebro. El axón de cada célula granular, eje troncal principal portador de la señal eferente, emerge verticalmente desde la capa granular, estableciendo múltiples puntos de entrada con la célula de Purkinje ubicada sobre ella. El axón se escinde después en dos segmentos que se extienden en direcciones opuestas. Estos segmentos se alinean y forman fibras paralelas que cursan a través de las dendritas de una célula de Purkinje, a modo de hilos del tendido eléctrico, con lo que proporcionan una misma señal a muchos cientos de células de Purkinje. Los neurocitos granulares se comunican también con otros tres tipos de neuronas: estrelladas, en cesto y células de Golgi. Ese trío contribuye a modular las señales emitidas tanto por las células granulares como por las de Purkinje. Se trata de una configuración básica en todos los cerebelos, lo que indica que ha de ser parte integrante de su función.



ALICE Y. CHEN

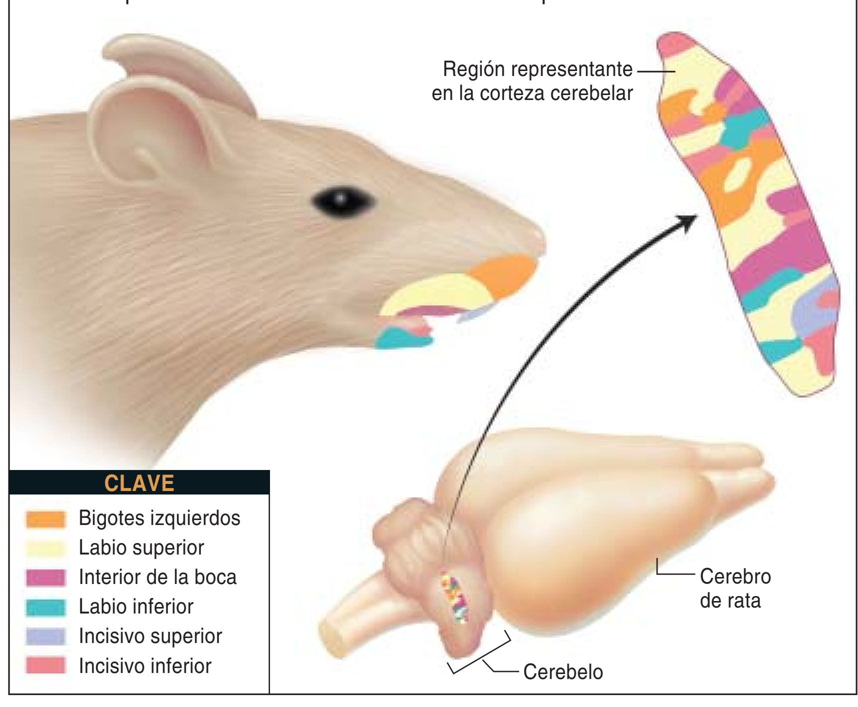
locidad y la dirección de figuras móviles. Por fin, Hermann Ackermann y sus colaboradores, también en Tubinga, observaron que los pacientes con degeneración cerebelar son menos diestros que los individuos sanos a la hora de discriminar palabras cuasihomófonas, como "rábido" y "rápido".

Las limitaciones que experimentan los lesionados cerebelares trascienden el lenguaje, la visión y la audición. Según Jeremy D. Schmahmann, del Hospital General de Massachusetts, los pacientes cerebelares, tanto infantiles como adultos, tienen dificultad para modular sus emociones: o bien no alcanzan a reaccionar, o bien reaccionan de forma excesiva ante estímulos que suscitan respuestas más moderadas en la mayoría de las personas. Otros investigadores han demostrado que los adultos con lesiones cerebelares son más lentos y más propensos a cometer errores en pruebas de razonamiento espacial; por ejemplo: determinar la concordancia de formas de objetos vistos desde distintos ángulos. Algunos autores han vinculado cerebelo y dislexia. Entre ellos, el grupo liderado por Rod I. Nicolson, de la Universidad de Sheffield; descubrió que los disléxicos y los pacientes de lesiones cerebelares presentan deficiencias similares en su capacidad de aprendizaje y que los disléxicos desarrollan menor actividad cerebelar al ejecutar ciertas tareas.

Otros estudios recientes sugieren la plausible intervención del cerebelo en la memoria a corto plazo, la atención, ciertas funciones mentales (planificación y organización temporal de tareas) y control de los impulsos. En este sentido, en 1992, Jordan Grafman y sus colaboradores, del norteamericano Instituto Nacional de Salud, observaron que los individuos con atrofia cerebelar tenían dificultad para planificar y organizar las etapas requeridas para resolver el problema de la Torre de Hanoi, en el cual es preciso ir colocando discos perforados de diferentes tamaños en una serie de vástagos, de acuerdo con determinadas reglas. A través de dos estudios independientes de formación de neuroimágenes, llevados a cabo en 1997, se reveló que el cerebelo de los voluntarios sanos se activaba cuando se les pedía que recordasen una lista de letras que habían oído recitar unos momentos antes, o cuando se les pedía que buscasen un motivo particular en una imagen concreta. En 2002, en otro ensayo mediante neuroimágenes, realizado por Xavier Castellanos, Judith L. Rapoport y otros, adscritos al Instituto Nacional de Salud Mental de EE.UU., se halló que el cerebelo de los niños con el trastorno de hiperactividad con déficit de atención —que se caracteriza por la falta de control de los estímulos— es pequeño. Por último, gracias a las investigaciones fundadas en formación de imágenes, se sabe que el cerebelo de personas y animales sanos se encuentra normalmente activo durante múltiples procesos sensoriales: la audición, la olfacción, la sed, la necesidad de alimento o de aire, la consciencia del movimiento del cuerpo y la percepción del dolor.

LA NATURALEZA FRACTURADA DEL CEREBELO

UN AREA PARTICULAR del rostro de una rata no queda representada en el cerebelo por una sola región. Resulta posible registrar actividad eléctrica en puntos diversos y muy distantes de la corteza cerebelar del múrido cuando se roza con una sonda el labio inferior del animal. Tal fragmentación puede permitir al cerebelo la integración de un abanico de información sensorial aferente obtenida por distintas partes del cuerpo durante el curso de una acción exploratoria.



ALICE Y. CHEN

Tocar, sentir

Nos contamos entre los investigadores que han llegado al convencimiento de que la teoría tradicional de la función cerebelar, que atribuye al cerebelo el control motor, no es adecuada para explicar los nuevos datos. Apoyamos nuestra opinión, primero, en el estudio exhaustivo de las regiones cerebelares activas durante el tacto. Uno de los autores (Bower) inició esa línea de trabajo hace más de 20 años, mientras preparaba la tesis doctoral en el laboratorio de Wallace I.

Los autores

JAMES M. BOWER y LAWRENCE M. PARSONS pertenecen ambos al Centro de Investigación de Formación de Imágenes del Centro de Ciencias de la Salud de Texas, en San Antonio, donde Bower es profesor de neurobiología computacional, y Parsons, de neurociencia cognitiva.

Welker, en la Universidad Wisconsin de Madison. Se valía de la microtopografía, técnica que le permitía registrar la actividad eléctrica de pequeñas asociaciones de neuronas de cerebros de ratas, cuando éstas eran levemente tocadas en distintas zonas del cuerpo.

Tales estímulos táctiles evocaban actividad a través de una extensa región del cerebelo. Lo que es más, el mapa ofrecía un aspecto fracturado, pues ocurría a menudo que regiones vecinas del cerebelo eran destino de señales procedentes de regiones muy dispares del cuerpo, al par que áreas que en el cuerpo se hallan alejadas quedaban representadas en el cerebelo una junto a otra. Tal sistema de cartografía difiere mucho del correspondiente a la corteza cerebral; en ésta, las relaciones espaciales entre áreas de la superficie del cuerpo se conservan en las regiones corticales que responden a las señales venidas de dichas áreas o envían señales a ellas.

Aunque la naturaleza fracturada de los mapas cerebelares sea insólita, todavía más sorprendente resultó el descubrimiento de que el cerebelo de la rata recibía primariamente señales procedentes del rostro del animal. En un principio, tal fenómeno provocó desconcierto, porque Snider había demostrado con anterioridad que la mayor parte de la región táctil del cerebelo del gato recibe señales procedentes de las patas anteriores y que, en los monos, el grueso de esa región se activa cuando se les tocan los dedos.

Un coordinador sensorial

Ante las diferencias entre las regiones de superficie corporal que están representadas en los cerebelos de animales distintos, la cuestión pasó a ser: ¿en qué se parecen la boca de una rata, las patas anteriores de un gato y los dedos de un mono? Los estudios de Wisconsin arribaron a la conclusión de que cada estructura es utilizada por cada animal para averiguar, a través del tacto, cómo es su entorno. Quien tenga un gato sabrá cuántas diabluras puede hacer con las patas anteriores; quien esté familiarizado con los niños se habrá dado cuenta de cuán activa —y a veces, dolorosamente— utilizan los pequeñuelos sus deditos para adquirir información sobre el mundo circundante. Pero a las ratas puede resultarles peligroso servirse de la boca. La estructura fracturada de los mapas del tacto en el cerebelo respaldaba la idea de que la región estaba, aunque no se supiera cómo, comparando los datos sensorios llegados de las distintas partes del cuerpo que cada animal utilizaba para explorar su mundo. Estos mapas parecían estar organizados de acuerdo con el uso que se le diera a cada parte del cuerpo; de ningún modo, según su proximidad absoluta en la superficie corporal. La idea de que el cerebelo de la rata comparaba la información sensorial procedente de diversos puntos de la cara recibió confirmación adicional de diversos modelos y experimentos que examinaban la respuesta del cerebelo ante estos impulsos. Nació de ello una nueva hipótesis sobre la función cerebelar: a tenor de la misma, el cerebelo se hallaba específicamente implicado en la coordinación de la adquisición de datos sensoriales por el cerebro.



RUDOLF VAN'T HOFF padece una lesión cerebelar provocada por atrofia espinocerebelar de tipo I. Se trata de una enfermedad genética, muy rara; acostumbra manifestarse a mediana edad. A nuestro hombre le ha afectado el equilibrio, el habla y la coordinación, amén de mermar su capacidad de discriminación entre ciertos sonidos. Cuando trabaja en su tractor, Van't Hoff se ayuda de un "pulpo" para estabilizar su cuerpo.

CHRIS HARTLOVE

No es difícil formular nuevas ideas sobre la función cerebral; muy distinto es lograr que tales ideas sean aceptadas en una especialidad que, ya a mediados del siglo XIX, había llegado a la conclusión de que el cerebelo constituía una estructura motora. En este caso, la tarea resultaba más difícil todavía, porque existe en el cerebro una vinculación muy íntima entre los sistemas sensorial y motor, especialmente en los que interviene el tacto. Para cerciorarnos de que estábamos observando solamente los efectos de la actividad sensorial, y no los de la motora, se imponía investigar en humanos; a diferencia de las ratas, las personas pueden obedecer instrucciones explícitas sobre cuándo han de moverse, y cuándo no. En ese momento dio comienzo la asociación de los autores del artículo. En colaboración con Peter T. Fox, del hospital clínico de la Universidad de Texas en San Antonio, pergeñamos un plan de trabajo. Fundábase en la obtención de imágenes de la actividad neurológica, con el fin de comparar la magnitud de la actividad cerebelar inducida cuando se solicitaba a los voluntarios que utilizasen los dedos en tareas de discriminación mediante el tacto, o se les pedía, sencillamente, que tomasen objetos pequeños y los soltasen. De acuerdo con todas las teorías precedentes sobre la función cerebelar, sería de prever que el control motor fino de los dedos, re-

UN SONDEO DE LA FUNCION CEREBELAR

AL OBJETO DE ACOTAR los posibles roles del cerebelo en la coordinación de los movimientos y en la integración de las señales sensoriales aferentes, proyectamos un experimento a desarrollar en cuatro fases. Nos apoyamos en la resonancia magnética funcional, técnica que nos revelaría qué actividad mostraba el cerebelo de seis personas sanas mientras estaban sintiendo un estímulo en los dedos, sin moverlos, o bien asiendo y soltando objetos pequeños. En un primer supuesto procedimos a inmovilizar las manos del sujeto y rozarle levemente las yemas de los dedos con papel de lija (*a*). A veces se les pedía que comparasen el grano de los dos tipos de lija (*b*). Se trataba en ambos casos de tareas estrictamente sensoriales, pero en la segunda cada persona tenía que discriminar qué era lo que percibía en sus manos.

El segundo supuesto entrañaba aspectos sensoriales y motores. El voluntario introducía las manos en sendas bolsas, que contenían bolitas de madera de distintas formas y texturas. En la primera tarea (*c*) se le pedía al sujeto que tomase bolas al azar y las soltase, sin prestar atención a sus formas. En la segunda, el voluntario debía comparar las formas y sensaciones táctiles de dos bolas, siempre que tomaba una en cada mano.

El cerebelo no dio apenas signos de actividad durante la tarea que solamente requería asir y soltar las bolas (*c*). Pero exhibía actividad máxima cuando los sujetos estaban evaluando sus sensaciones, tanto en movimiento (*d*) como sin él (*b*). Estos y otros hallazgos abonan nuestra hipótesis: la actividad principal del cerebelo corresponde al procesamiento de información sensorial, no al control del movimiento.

SENSACION PASIVA

Sin movimiento

a

Papel de lija fina

Papel de lija gruesa

Cerebelo

Area activa

Con movimiento

c

COMPARACION ACTIVA DE SENSACIONES

Sin movimiento

b

Con movimiento

d

querido para asir y dejar caer los objetos pequeños, provocase una gran actividad en la región cerebelar asociada con el tacto. Encontramos, por el contrario, muy poca actividad cerebelar en dicha región durante la tarea de asir y soltar, mientras que la exploración táctil, palpando con los dedos, suscitaba una intensa respuesta cerebelar. Esta observación vino en apoyo de nuestra idea de que el cerebelo se encuentra más implicado en la función sensorial que en la sólo motora y, en particular, que se mantiene sumamente activo durante el proceso de adquisición de datos sensoriales.

La vida sin cerebelo

Nuestra hipótesis de adquisición sensorial no es sino una de las varias teorías nuevas que han surgido ante las pruebas, crecientes, de que el cerebelo no sólo interviene en el mero control motor. En muchos casos, ha sido posible dar acomodo a los nuevos datos sin más que ampliar las teorías motoras existentes para dar cuenta de resultados no motores. Ivry, por ejemplo, ha apelado a una hipótesis de "temporización generalizada" de la función cerebelar; sugiere ésta que el cerebelo controla la secuencia temporal de los movimientos corporales (así, la coordinación de los cambios en los ángulos de las articulaciones) para permitir a los individuos estimar la duración de señales sensoriales, como las proporcionadas por la vista o el oído.

Otros investigadores han postulado que el cerebelo no sólo facilita los movimientos finos, sino que también "alisa" el procesamiento de información relacionada con el estado de ánimo y el pensamiento. Schmahmann expresó ya esta opinión en 1991; en 1996, Nancy C. Andreason, de la Universidad de

MATT COLLINS (*ilustraciones*); REIMPRESION DE JIA-HONG GAO ET AL. EN *SCIENCE*, VOL. 272; 26 DE ABRIL, 1996 (*escáneres cerebrales*)

Iowa, adaptó a la esquizofrenia esta hipótesis. Andreason sostiene que la causa subyacente a los trastornos mentales característicos de esa patología podrían ser deficiencias cerebelares. Otros han propuesto que las regiones de cerebelo que han experimentado un desarrollo más impresionante a lo largo de la evolución humana proporcionan respaldo computacional para aquellas tareas psicológicas que resulte posible descargar de la corteza cuando ésta sufre sobrecarga.

Conforme crecía el número de afecciones que conllevan cambios en la actividad cerebelar, se han ido atribuyendo cada vez más funciones al cerebelo. Pero todavía está por explicar de qué modo puede una estructura cuya circuitería neuronal está organizada según un patrón simple y repetitivo desempeñar una función tan integral en funciones y conductas muy dispares.

Y lo que resulta incluso más desconcertante: tras una lesión cerebelar cabe la recuperación. Aunque la extirpación completa del cerebelo destruye la coordinación de los movimientos, los individuos (en especial, los jóvenes) pueden, con suficiente tiempo, recuperar en grado considerable la función normal. Aunque tal plasticidad constituye una característica general del cerebro, lesiones similares de las regiones sensoriales primarias o de control motor en la corteza cerebral provocan, lo mismo en animales que en humanos, un deterioro grave y permanente de funciones específicas.

Ante semejante capacidad de recuperación tras la extirpación del cerebelo, algunos declaran, medio en serio medio en broma, que su función bien pudiera estribar en compensar su propia ausencia. Es, sin embargo, sumamente improbable que una estructura tan grande e intrincada como el cerebelo carezca de función o que se trate de una estructura vestigial. La función cerebelar parece, en cambio, permitir al resto del cerebro la compensación de su ausencia en un grado considerable.

Algunas teorías cerebelares, incluidas las basadas en el control motor, sí han proporcionado explicación a la enigmática capacidad de recuperación. A nuestro juicio, la facultad del cerebro para compensar la ausencia de cerebelo entraña una sutil función de respaldo general. Si se admite la hipótesis de coordinación sensorial que sostenemos los autores, el cerebelo no es responsable directo de ninguna conducta ni proceso psicológico. Funcionaría como una estructura de respaldo o apoyo para el resto del cerebro. Tal respaldo entraña la supervisión de los datos sensoriales aferentes y la incesante realización de ajustes muy finos sobre la forma en que es adquirida dicha información; su objetivo consistiría en asegurar la máxima calidad posible de las señales que proceden de los sentidos.

En nuestra opinión, estos ajustes adoptan la forma de cambios extraordinariamente sutiles en las posiciones de los dedos humanos al palpar, de las barbas de la rata, en la retina o en el oído interno. Sería de esperar que el cerebelo, dada su estructura de apoyo, tuviera cierto nivel de actividad en múltiples situaciones, en especial, en las que requieren un control cuidadoso de los datos sensoriales aferentes e, incluso, de los recordados. Otros sistemas cerebrales pueden, de ordinario, compensar la falta de coordinación de los datos sensoriales mediante diversas tácticas de procesamiento, en el caso de que el cerebelo resulte lesionado o extirpado.

De hecho, los estudios de coordinación motora sugieren que quienes padecen lesiones cerebelares tienden a lentificar y simplificar sus movimientos, estrategias que son razonables para compensar la carencia de información sensorial de alta calidad. Una generalización interesante y de importancia de esta idea concierne al funcionamiento continuado de un cerebelo defectuoso, que tendría consecuencias más graves que su remoción completa. Aunque otras estructuras cerebrales puedan compensar la falta absoluta de control de datos sensoriales, sería de esperar que un control sistemáticamente defectuoso provocase una disfunción continua en otras regiones del cerebro que intentasen valerse de datos erróneos. Este tipo de efecto pudiera explicar las recientes conjeturas relativas a la participación cerebelar en el autismo, ejemplo típico de trastorno por el que los enfermos son incapaces de responder a los datos que les llegan desde los sentidos.

Hipótesis como la nuestra conllevan un útil recordatorio para la investigación futura: la presencia de actividad en una región del cerebro no asegura que esté participando directamente en una determinada conducta o proceso psicológico. La mayor parte de la maquinaria oculta bajo el capó de un automóvil tiene la función de contribuir al funcionamiento del motor. Podríamos formular todo tipo de hipótesis sobre el papel del radiador para la propulsión —por ejemplo, estableciendo una correlación entre el aumento de temperatura con la distancia recorrida, o bien observando que el coche deja de funcionar si se elimina el radiador—, pero el radiador no es el motor.

Si el cerebelo constituye primordialmente una estructura de apoyo, no contribuirá de una forma directa a la coordinación motora, ni a la memoria, la percepción, la atención o el razonamiento espacial, ni a cualquiera de las funciones que se le han atribuido en los últimos años. Aunque esta teoría sea una de varias en contienda para explicar los nuevos y sorprendentes hallazgos en torno al cerebelo, parece incuestionable que nuestra forma de pensar acerca de dicha estructura —y en consecuencia, de nuestra concepción del cerebro— está a punto de cambiar.

Bibliografía complementaria

THE CEREBELLUM AND COGNITION. Selección de Jeremy D. Schmahmann. Academic Press, 1997.

CEREBELLAR CONTRIBUTIONS TO COGNITION AND IMAGERY. Richard B. Ivry y Julie A. Fiez, en *New Cognitive Neurosciences*. Selección de Michael S. Gazzaniga. MIT Press, 2000.

THE CEREBELLUM: RECENT DEVELOPMENTS IN CEREBELLAR RESEARCH. Selección de Stephen M. Highstein y W. Thomas Thatch. New York Academy of Sciences, 2002.

CIENCIA Y SOCIEDAD

Información cuántica

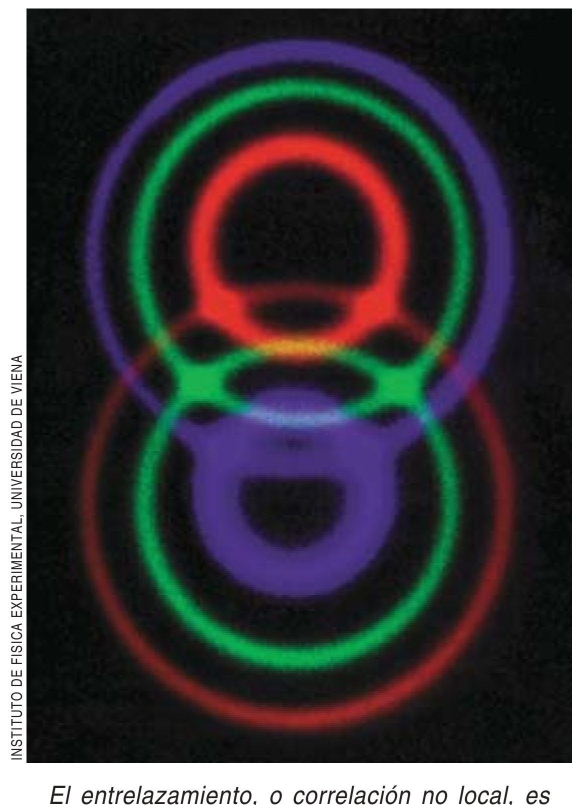
Estado de la cuestión

La información cuántica es un área de investigación en rápido desarrollo que comprende distintas disciplinas científicas, entre las que se incluyen la física, las ciencias de la computación y la matemática aplicada. En sentido amplio podríamos decir que es el resultado de la confluencia de cuatro campos diferentes del saber: la teoría cuántica (física), la teoría (clásica) de la información, la teoría de la computación y la criptografía.

En la teoría cuántica, desarrollada en el primer tercio del siglo pasado para entender la estructura atómica, se cimientan los desarrollos tecnológicos de mayor éxito de la segunda mitad del siglo XX. Es la base y el futuro de la química moderna, porque fundamenta con rigor el enlace químico; de los sólidos, puesto que, al explicar la interacción entre los electrones y los iones de la red cristalina, la teoría de bandas permite entender el comportamiento de los conductores, aislantes y semiconductores, presentes en toda la tecnología electrónica actual; de los láseres; de las microscopías electrónica y de efecto túnel; de la superfluidez; de la superconductividad, etc. Es difícil encontrar un producto cotidiano que utilice técnicas actuales y no tenga de un modo u otro su fundamento en la teoría cuántica, hasta el punto de que se estima que el 25 % del PIB mundial depende de un modo u otro de ella.

La necesidad de abordar los problemas de la información, la computación y la criptografía, aspectos básicos de las sociedades modernas, desde un punto de vista cuántico tiene una doble motivación. Por un lado es bien conocido que el número de transistores por unidad de superficie en un microchip se dobla, a costo constante, cada 18 meses, aproximadamente (ley de Moore), por lo que nuestros ordenadores parecen lentos o quizás inútiles a los tres o cuatro años de haberlos comprado. Esta reducción del tamaño tiene su límite en el mundo microscópico, e incluso antes, en el mesoscópico. El horizonte temporal de ese límite se sitúa hoy día en unos 30 años. Cuando se llegue a ese dominio, las leyes de la física que rigen el comportamiento de los componentes de los ordenadores cambiarán, pasando de la física clásica a la cuántica. Desarrollar las bases teóricas del computador cuántico, que utilizaría las propiedades cuánticas genuinas para funcionar, es casi una necesidad de nuestra sociedad hoy. Pero no es ésta la única motivación. Por otro lado, la mecánica cuántica permite hacer tareas y resolver problemas que no tienen solución, que son imposibles, en las teorías (clásicas) de la información, la computación y la criptografía, y que serán esenciales para el desarrollo futuro de nuestras complejas sociedades: así, la posibilidad de computar con muchísima más rapidez o la de transmitir de modo absolutamente secreto flujos de información. La información cuántica le debe el ser mucho más poderosa que la información clásica a la existencia de estados entrelazados, o que exhiben correlaciones estadísticas aun cuando no hay un nexo causal, o transferencia de información, entre ellos (principio de superposición y composición tensorial de estados de varias partículas).

Algunos momentos esenciales en la constitución de la información cuántica fueron, primero, los experimentos que contrastaron las desigualdades de Bell, relaciones numéricas ligadas al entrelazamiento, que establecieron de modo inequívoco la existencia de éste; luego, la construcción del protocolo de teleportación y su posterior verificación experimental, y la introducción, por Peter Shor, de un algoritmo cuántico que permitiría resolver problemas imposibles de superar con los ordenadores convencionales.



INSTITUTO DE FISICA EXPERIMENTAL, UNIVERSIDAD DE VIENA

El entrelazamiento, o correlación no local, es esencial en la información cuántica. Se pueden producir fotones entrelazados cuando un haz láser ultravioleta incide sobre un cristal no lineal dando lugar a un proceso denominado "conversión paramétrica a la baja". Existe una pequeña probabilidad de que uno de los fotones incidentes dé lugar a un par de ellos de mayor longitud de onda (el conjunto de anillos concéntricos de las figuras). Bajo determinadas condiciones unos de los fotones tendrá polarización horizontal y el otro vertical. Si se superponen los haces se forma un estado entrelazado, como los que se utilizan en información cuántica

Como ya se ha indicado, las tres aplicaciones más importantes en esta área son la teleportación, la criptografía y la computación cuánticas.

Una línea resumiría qué se está haciendo hoy en el campo de la información cuántica: los teóricos estudian situaciones donde intervienen estados cuánticos coherentes (es decir, con entrelazamiento), y los experimentadores tratan de manejarlas, de momento con un número reducido de elementos. He aquí algunos de los problemas abiertos que se están estudiando. Por el lado teórico:

I) La búsqueda de criterios generales que determinen cuándo un estado es separable en componentes no entrelazados y cuándo no.
II) La cuantificación del entrelazamiento para estados de varias partículas y determinación de la equivalencia entre sus distintas medidas.
III) La "destilación" del entrelazamiento en estados mezcla (mezclas estadísticas ordinarias de estados cuánticos). Estudio de los estados entrelazados no destilables.
IV) El desarrollo de nuevos algoritmos cuánticos.
V) Las capacidades de los canales y los problemas de superaditividad (magnitudes de un estado mayores o iguales que la suma de esas mismas magnitudes para los componentes del estado).

Desde el punto de vista experimental, el grupo de Nicolas Gisin, en Ginebra, y el de Richard Hughes, en Los Alamos, están realizando progresos muy notables en criptografía cuántica. Están enviando mensajes que no pueden ser descifrados, o que, si son descifrados, quienes los intercambian pueden saber con total seguridad si alguien los ha interceptado o no, el primero mediante la fibra óptica, con un alcance de 67 kilómetros, y el segundo a través de la atmósfera, incluso cuando el emisor se halla a 10 kilómetros del receptor del mensaje. Se están también explorando los sistemas físicos que, con la deseada propiedad de la escalabilidad (capacidad de adaptarse a tareas de distinta magnitud), puedan resultar más adecuados para procesar la información cuántica. Citemos entre ellos los sistemas ópticos cuánticos (cavidades resonantes de alta calidad e iones confinados en trampas electromagnéticas) y la creciente actividad en el campo del estado sólido, donde se están realizando puertas lógicas por medio de puntos cuánticos y de uniones Josephson superconductoras (realizadas ya para un qubit —la unidad cuántica de información—, en ensayo para dos). Otros desarrollos experimentales recientes con incidencia en este campo son la generación de impulsos individuales de fotones (en el Instituto Max Planck); el desarrollo de detectores de fotones con una eficiencia de casi el 100 % (propuestos por Kwiat y por Imamoglu); la sincronización de relojes y sistemas de posicionamiento valiéndose del entrelazamiento (en el Instituto de Tecnología de Massachusetts); demostraciones experimentales de la corrección de errores (en el Instituto Nacional de Patrones y Tecnología de Estados Unidos); la clonación con alta fidelidad de sistemas cuánticos individuales (Oxford); etc.

Además de todas estas excitantes perspectivas, hay otros dos aspectos en la información cuántica que se deben mencionar, a saber: el valor añadido que tendría el desarrollo de nuevas áreas técnicas en una futura sociedad de la información, hacia la que caminamos, y su contribución, en nuestro caso, a los objetivos sociales de Europa. Más que desplazar a la actual tecnología de la información, la nueva información cuántica se haría un lugar a su lado, en la ejecución de tareas imposibles con la técnica actual. A nadie se le oculta la trascendencia económica de, por ejemplo, transmitir de modo absolutamente seguro flujos de información (piénsese en Internet o en la necesaria codificación de un mensaje cuando se efectúa una transacción económica con la tarjeta de crédito), y así lo han entendido algunas de las empresas más importantes del mundo, como IBM, Hewlett Packard, Laboratorios Bell, Microsoft, etc., e instituciones y agencias gubernamentales de los países mas desarrollados, fundamentalmente en EE.UU. (la Universidad de Harvard, el Caltech, el MIT, la NASA, el Ministerio de Defensa, etc.), en la Unión Europea, a través de la Fundación Europea de la Ciencia, y en distintos países europeos, en particular el Reino Unido, Austria y Alemania.

MIGUEL FERRERO
Departamento de Física
Universidad de Oviedo

Genoma humano

Orden interno

El genoma humano contiene aproximadamente 32.000 genes. Estos determinan la composición de las proteínas que regulan el desarrollo, metabolismo y demás funciones del organismo. Todos los genes del núcleo se hallan distribuidos entre los 23 cromosomas que constituyen el genoma humano. En conjunto, los genes ocupan menos del 5 % del genoma; si se repartieran de manera uniforme, cada gen debería estar separado de los demás. No ocurre tal. La densidad génica varía a lo largo de los cromosomas. Hay extensas regiones casi desiertas, mientras que en otras los genes, muy próximos, se superponen.

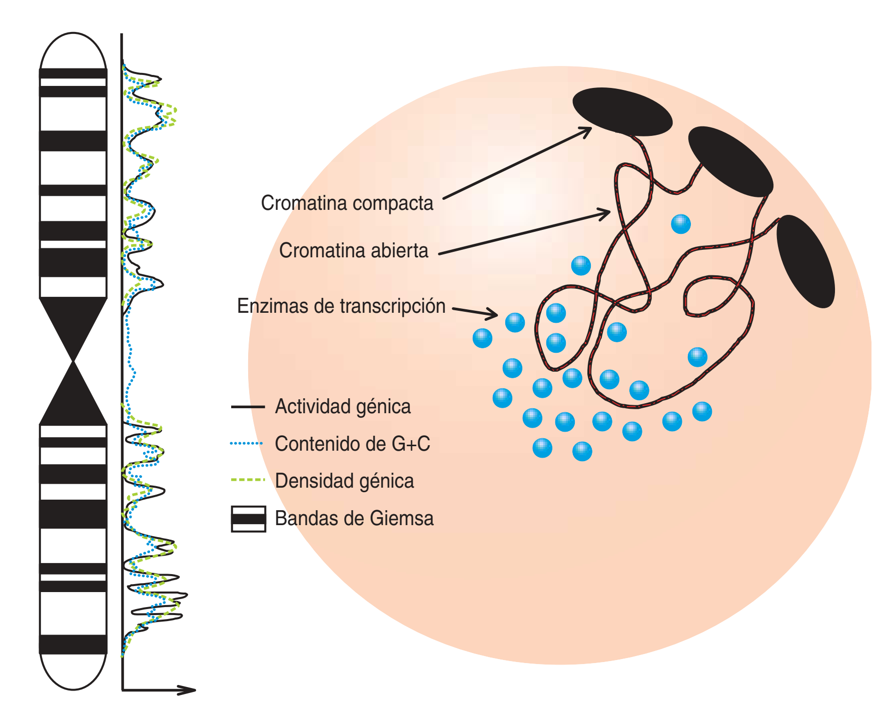
¿Qué determina la posición y el orden de los genes en el genoma? Suele darse por supuesto que las presiones selectivas sobre dicha ubicación son mínimas. Ahora bien, ello comportaría que la posición y orden actual de los genes correspondiesen a una distribución aleatoria. Sin embargo, los análisis realizados han revelado patrones, difíciles de explicar por simple azar. Tales pautas sugieren que la distribución de genes se encuentra vinculada a la compleja estructura de los cromosomas.

Los cromosomas se dividen en regiones que difieren en estructura y composición de bases. De las cuatro bases nitrogenadas del ADN, la proporción de guanina y citosina, con respecto a adenina y timina, puede

ser de hasta un 60 % en algunas regiones del genoma; en otras alcanza sólo un 35 %. La densidad de genes se halla asociada a estas variaciones en composición de nucleótidos. Los genes tienden a concentrarse en regiones ricas en guanina y citosina (G + C).

Desde finales del siglo XIX, se han venido identificando, con técnicas de coloración Giemsa aplicadas a células en mitosis, bandas de intensidad cromática variable y transversales a lo largo de los cromosomas. Recientemente, Ferreira, Carmo Fonseca y colaboradores, de la Universidad de Lisboa, observaron que, durante la división celular, las bandas pálidas se replicaban antes que las obscuras. Las bandas obscuras, más compactas también, tienden a ocupar regiones hacia la periferia del núcleo. Se ha descubierto que las variaciones en bases nitrogenadas están relacionadas con las bandas de Giemsa: las bandas claras y menos compactas contienen una mayor proporción de G + C, mientras que las obscuras y compactas son más ricas en adenina y timina.

¿Cuál es el papel de la estructura cromosómica para el funcionamiento génico? Se conoce desde hace tiempo una relación entre la densidad de los genes y características estructurales de los cromosomas. Más de un 80 % de los genes se encuentran en las bandas de Giemsa más claras. Sin embargo, las razones de esta concentración no son conocidas.



Puntos principales del modelo propuesto de orden génico y su relación con la estructura cromosómica. Algunos de los puntos no han recibido todavía confirmación. Los niveles de actividad promedio de los genes en cada región se hallan en correlación con la densidad génica, el contenido de guanina y citosina y las bandas de Giemsa. Las bandas de Giemsa se distribuyen desigualmente dentro del núcleo. Las bandas más obscuras y compactas se sitúan en la periferia, mientras que las más claras y ricas en genes se encuentran en el interior del núcleo donde el acceso a las enzimas de transcripción es mayor

Los genes varían en sus niveles de actividad, ya que no se requieren todas las proteínas en todos los tipos celulares, ni en la misma cuantía. En una muestra de 10.000 genes, el equipo encabezado por Lercher, de la Universidad de Bath, observó que los genes se agrupaban según el número de tejidos en los que se expresan; dicho de otro modo: hay genes que se expresan en numerosos tejidos y genes que se expresan en pocos tejidos; ambos tienden a estar cerca de los genes de expresión similar. Esta fue la primera evidencia sistemática de la importancia de la posición en el cromosoma para la actividad génica.

El modelo de *compartamentación funcional* del genoma aporta un marco común para explicar las variaciones de densidad y orden de los genes observados a lo largo de los cromosomas. Este propone que hay regiones más favorables que otras para la actividad de los genes. A tenor del modelo, la estructura de una región cromosómica particular constituye un factor determinante para el potencial de actividad de los genes. Las regiones correspondientes a bandas de Giemsa claras y menos compactas se sitúan en la periferia del núcleo, donde pueden acceder con mayor facilidad los complejos de transcripción. Ahí era donde cabía esperar la mayor densidad de genes, lo que se ha comprobado.

Christoph Cremer, de la Universidad Ludwig Maximilians, y Thomas Cremer, de la Universidad de Ruprecht, establecieron la hipótesis según la cual los genes que se expresan en más tipos celulares, deberían estar concentrados donde la actividad de transcripción fuera más intensa. Los datos del grupo de Lercher abonan ese planteamiento. Se ha observado que los genes que se expresan en un número mayor de tejidos se asientan en regiones ricas en G + C. Si el modelo de Cremer y Cremer es correcto, podemos esperar que los genes de amplia expresión se localicen en las bandas de Giemsa claras y menos compactas de los cromosomas.

De confirmarse tal ubicación, la idea del genoma como un depósito pasivo de genes debería ceder paso a un modelo en el que la estructura de los cromosomas desempeña una función activa en la determinación de los patrones generales de expresión y en la evolución de las especies. Pensemos, por ejemplo, en mutaciones que comportan la eliminación o duplicación de segmentos de cromosoma: las consecuencias para el organismo podrían derivarse no sólo de la ausencia de los genes eliminados o añadidos, sino también de los cambios en la reorganización cromosómica nuclear.

Las observaciones anteriores sugieren que los factores relacionados con la estructura cromosómica intervienen en la determinación del orden de los genes en eucariotas, en particular en genomas de mamíferos. No ocurre así en los genomas de organismos procariotas (las bacterias, por ejemplo): aquí es normal que los genes implicados en un mismo proceso celular se encuentren agrupados en operones y su actividad se regule de forma conjunta. Carecemos de pruebas de que, en los insectos y los mamíferos, por citar dos ejemplos de genoma de eucariota, la corregulación sea un determinante importante en la localización para la mayoría de los genes.

ARAXI URRUTIA ODABACHIAN
Dpto. Biología y Bioquímica
Universidad de Bath

Agua

Estructura vibracional

El agua se encuentra entre las sustancias más comunes de nuestra vida cotidiana. Ocupa más de las dos terceras partes de la superficie de nuestro planeta y casi igual proporción en el cuerpo humano. El agua y las soluciones acuosas desempeñan una función esencial en los procesos fisicoquímicos que ocurren en el planeta y en los procesos biológicos. No se concibe, por ahora, forma posible de vida sin agua.

Su aparente trivialidad y que su molécula sea una de las más simples (H_2O) esconden la realidad: el agua es un líquido muy complejo y muestra muchas propiedades anómalas tanto microscópicas como macroscópicas.

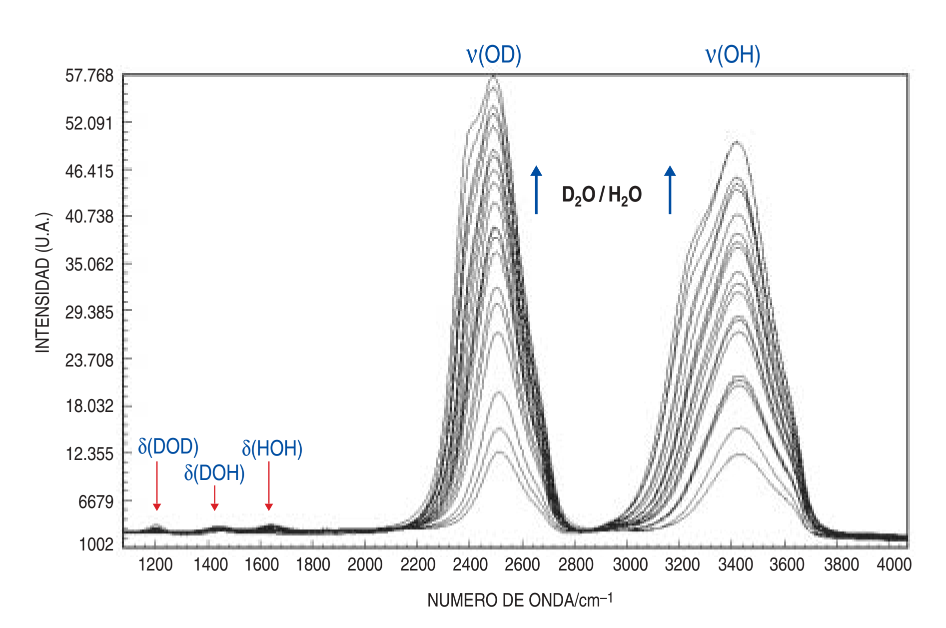
Unas son muy conocidas, como la de que alcance un máximo de densidad a 4 ºC, en la fase líquida, de forma que la fase sólida, el hielo, es menos densa que la líquida; tiene una enorme importancia en los ciclos climáticos y en el comportamiento de la biosfera. Otras, las fases metaestables, supercongelada y supercalentada, en las que existen fenómenos críticos, no resultan tan conocidas. Y algunas desafían el sentido común, como el efecto Mpemba, denominado así en honor del tanzano Erasto Mpemba, quien observó que el helado se haría antes si la mezcla estaba caliente. Para la mayor parte de estas propiedades del agua no existe una explicación única, por el momento. De todas formas, se reconoce que la clave radica en la estructura particular del agua, consecuencia del enlace entre los átomos de hidrógeno de unas moléculas H_2O y otras.

El enlace de hidrógeno confiere a las moléculas H_2O una geometría local semejante a la del hielo. Pero, a las temperaturas a las que permanece líquida el agua, la energía del enlace de hidrógeno no difiere mucho de la energía de agitación térmica. Por eso, la estructura del agua está fluctuando continuamente, con un período de fluctuación aproximado de 10^{-11} s. El entramado de enlaces de hidrógeno y su dinámica determinan las propiedades del agua.

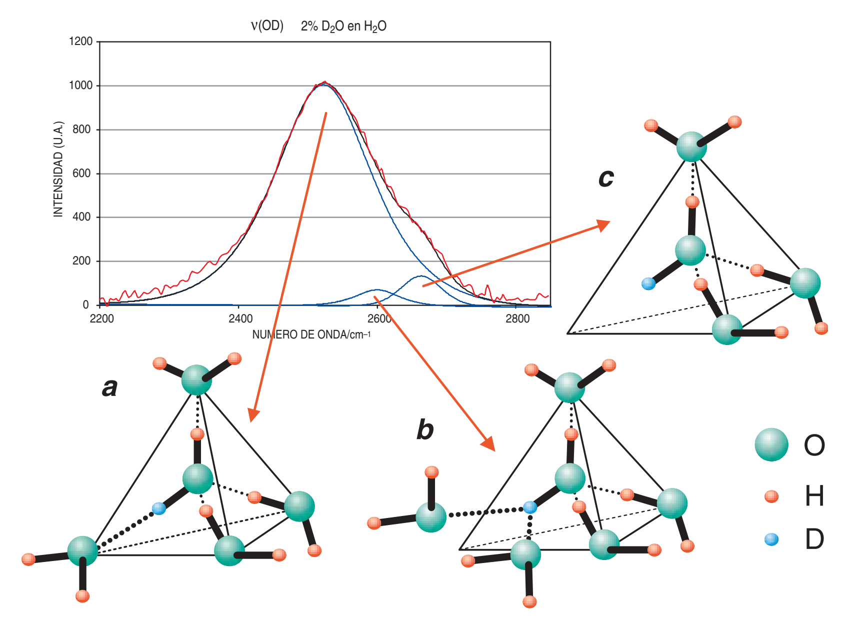
Una de las herramientas más eficaces para analizar la estructura dinámica del agua líquida es la espectroscopía vibracional, tanto en el modo de absorción de la radiación infrarroja (IR) como en el de dispersión inelástica de la luz, o espectroscopía Raman. El espectro Raman es además muy sensible a los efectos de corto alcance; resulta por ello útil en el estudio de la estructura local.

La dinámica vibracional del agua se puede representar de manera sencilla por medio de los modos de vibración internos y externos de las moléculas de H_2O en el seno de los enlaces de hidrógeno que forman con sus vecinas.

En estado libre, la molécula de H_2O tiene tres vibraciones internas fundamentales (el número de vibraciones de una molécula no lineal de N átomos es 3N-6), pero mientras el modo ν_2 de deformación de los enlaces OH se puede identificar fácilmente en el espectro experimental a 1640 cm^{-1}, no sucede lo mismo para los modos ν_1 y ν_3, que se extienden en la zona espectral 2800-3800 cm^{-1}.



1. Espectro Raman de la mezcla H_2O/D_2O a temperatura ambiente. Se observa la evolución de las bandas de tensión OD (~2500 cm^{-1}) y OH (~3450 cm^{-1}) así como las de deformación HOH (~1640 cm^{-1}), DOD (~1200 cm^{-1}) y HOD (~1450 cm^{-1}). A bajas concentraciones de cada especie isotópica se aprecia claramente la disminución de los acoplamientos intra e intermoleculares que complican sobremanera el espectro del agua líquida



2. Banda de vibración OD en una mezcla diluida al 2 % de D_2O en H_2O. Se observan tres componentes, que pueden relacionarse con tres situaciones estructurales distintas del oscilador OD: a) *coordinación tetraédrica de enlaces de hidrógeno; todos los enlaces son lineales o casi lineales.* b) *enlaces bifurcados; aquí se representa una posibilidad en la que la coordinación alrededor de una molécula dada es 5 en lugar de 4; otras configuraciones son también posibles.* c) *oscilador OD que no participa directamente en un enlace de hidrógeno*

La mayor parte de los modelos propuestos históricamente para explicar estas bandas de vibración parten de la idea, expuesta por primera vez por Bernal y Fowler en 1933, de que existe una configuración local de enlaces de hidrógeno que adopta, alrededor de cada molécula de agua, una geometría de tetraedro. Pero difieren en admitir o no la existencia de unidades química y estructuralmente diferenciadas en el líquido. A los que la admiten se los conoce como modelos de mezclas; a los que no, como modelos continuos.

El modelo de mezclas establece básicamente la existencia de dos tipos de osciladores OH, los enlazados por enlace de hidrógeno y los no enlazados. Viene a equivaler a admitir que hay moléculas de agua parcialmente libres en el seno del líquido.

Aunque la controversia entre ambos modelos se ha mantenido a lo largo de más de 50 años, en la última década se observa, junto a un renovado interés por el tema, nuevos resultados basados en la existencia de agregados bien localizados y estructurados: pentámeros, decámeros o agregados superiores, icosaédricos por ejemplo.

La posible existencia de los agregados icosaédricos resulta particularmente atractiva. Significaría que esa geometría debe de ser muy general y trascendería las estructuras sólidas no periódicas tanto de los cuasicristales como de fullerenos y agregados de gases nobles. Pero además nos recordaría a Platón, que asociaba los "sólidos platónicos" con los elementos esenciales del universo: "Demos a la tierra el cubo, al fuego la pirámide, al aire el octaedro y al *agua el icosaedro*".

La dificultad mayor para interpretar el espectro vibracional del agua se debe a la presencia de múltiples acoplamientos entre los osciladores OH, que complican sobremanera el espectro. La búsqueda de resultados experimentales que permitan arrojar nueva luz sobre la estructura dinámica del agua, y de paso dilucidar entre los modelos citados, depende, sin duda, de que se elimine en lo posible el efecto de los acoplamientos. Puede lograrse de varias maneras: disolviendo sales, variando la temperatura o con mezclas isotópicas H_2O/D_2O (donde D es el deuterio).

En una mezcla muy diluida de agua pesada (D_2O) en agua ligera (H_2O), la mayor parte de las moléculas toman la forma HOD. Los osciladores OD de estas moléculas se encuentran entonces casi totalmente desacoplados del entorno, constituido sobre todo por osciladores OH.

Las vibraciones de estos osciladores dependen sólo de la situación local en la que se encuentren. Si ésta consiste en una distribución continua de enlaces de hidrógeno, encontraremos una sola banda. Si el entorno contiene situaciones estructurales distinguibles, observaremos más de una banda.

El análisis detallado del perfil de las bandas OH y OD en condiciones diluidas muestra tres componentes (*figura 2*). Las de alta y baja frecuencia encajan perfectamente con la idea del modelo de mezclas a dos fases: los osciladores enlazados y los osciladores casi libres. La componente central ha sido relacionada por nosotros con la existencia de enlaces de hidrógeno bifurcados.

La posibilidad de que haya enlaces bifurcados abre interesantes perspectivas en la comprensión de la dinámica vibracional del agua. Por un lado, justifica en gran medida la existencia de los enlaces casi libres (un enlace OH bifurcado ocupa dos sitios aceptadores del protón), y por otro sirve a la formación e intercambio entre agregados, pues se la considera una vía de baja energía para hacer y deshacer el enlace de hidrógeno.

Aunque la utilidad de estos nuevos resultados para la interpretación de las propiedades físicas del agua está aún por ver, no cabe duda de que se los puede emplear para confeccionar modelos más elaborados y precisos de su estructura.

FERNANDO RULL PÉREZ
Departamento de Cristalografía y Mineralogía
Facultad de Ciencias, Universidad de Valladolid
Investigador del Centro de Astrobiología CSIC-INTA
Madrid

Josep-Maria Gili, Sergi Rossi, Beatriz Garcigoy y Jordi Grinyó

Veinticinco años después



1. Fotografía de la pared de "la pota del llop" tomada en el año 1978 a unos 30 m de profundidad. Se aprecia una gran densidad de gorgonias y otros invertebrados que recubren la pared

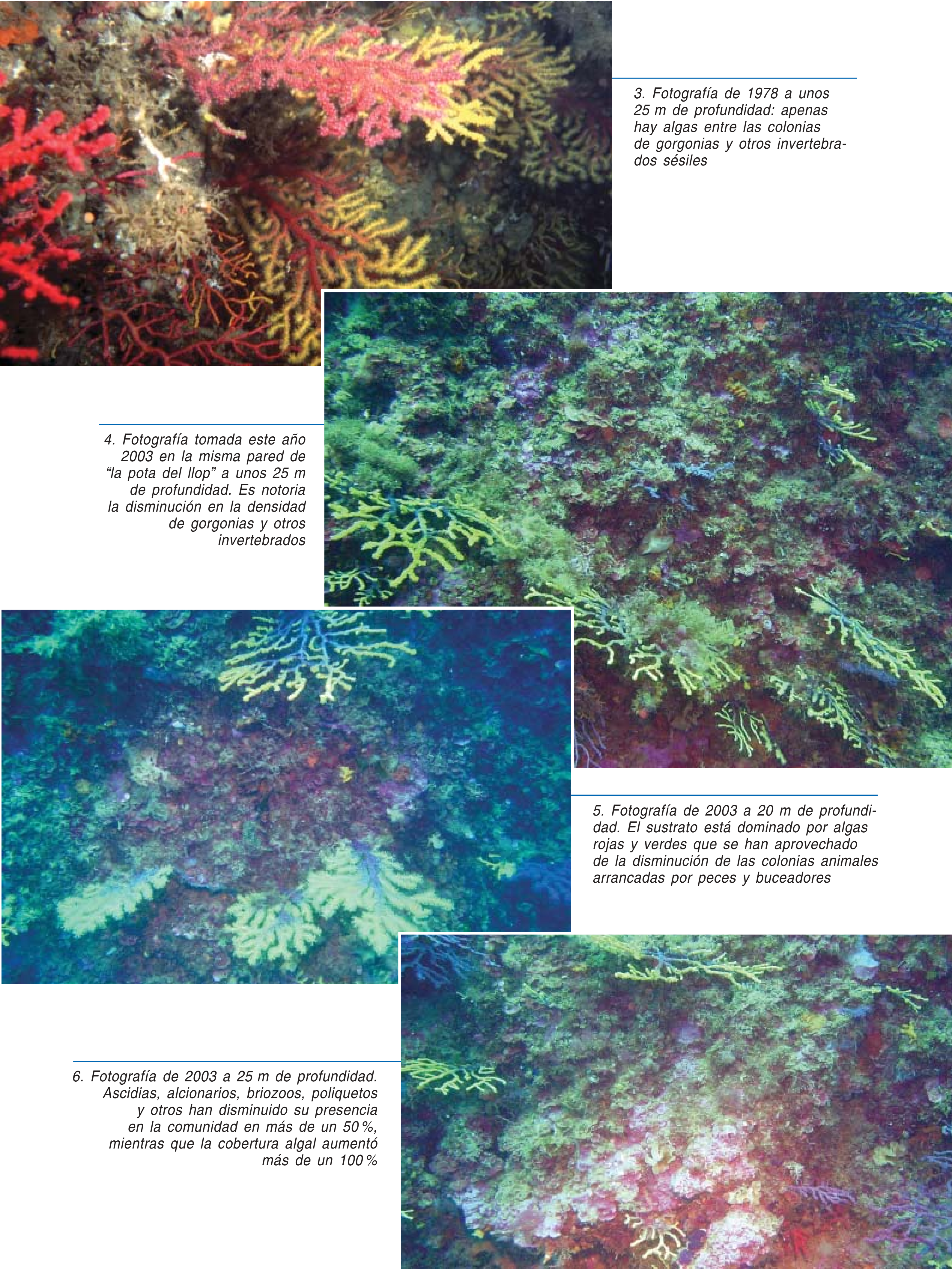


2. Fotografía de 1978 a unos 20 m de profundidad en la que se puede observar la zona de inicio (hacia mayor profundidad) de la comunidad de gorgonias, que venía precedida de alcionarios y briozoos. La población de los primeros ha mermado con la depredación de peces, cuyas poblaciones han incrementado espectacularmente en el parque. La caída de briozoos obedece, en buena medida, a la torpeza de algunos buceadores

Durante las últimas dos décadas las posibilidades de observar los fondos submarinos del Mediterráneo se han incrementado enormemente. Las nuevas técnicas y equipos puestos al servicio de los adictos a la inmersión han contribuido a que este deporte sea practicado por un número creciente de personas. La consecuencia positiva del hecho es que se ha extendido el conocimiento de los fondos marinos. Pero los efectos negativos no han tardado en aparecer.

El buceo frecuente en zonas determinadas ha terminado por perturbar, de modo muy señalado, a los organismos sésiles. En los ecosistemas terrestres el peligro de degradación antropogénica se ha podido paliar en parte con la creación de parques naturales. Los ecosistemas marinos costeros cumplen también la función importante de proteger el medio de la degradación. Sin embargo, en uno de los lugares más extraordinariamente ricos y diversos del Mediterráneo como son las islas Medas ha ocurrido un efecto totalmente contrario.

El archipiélago constituye una reserva natural marina desde 1990 y reserva de pesca desde 1983. No obstante tales restricciones legales, la instrumentalización financiera asociada al turismo en sacrificio de su alto valor ecológico a proteger ha traído la degradación evidente de una parte de sus fondos marinos. La visita permitida de unos 2000 buceadores diarios en las aguas de un archipiélago de 511 hectáreas, de las que apenas 93 son de estricta protección, ha causado que los ricos fondos de gorgonias y otros invertebrados bentónicos presenten hoy en día un aspecto muy distinto del que tenían hace 25 años. Tal alteración se refleja, de una forma dramática, en la llamada "pota del llop" (pata del lobo).



3. Fotografía de 1978 a unos 25 m de profundidad: apenas hay algas entre las colonias de gorgonias y otros invertebrados sésiles

4. Fotografía tomada este año 2003 en la misma pared de "la pota del llop" a unos 25 m de profundidad. Es notoria la disminución en la densidad de gorgonias y otros invertebrados

5. Fotografía de 2003 a 20 m de profundidad. El sustrato está dominado por algas rojas y verdes que se han aprovechado de la disminución de las colonias animales arrancadas por peces y buceadores

6. Fotografía de 2003 a 25 m de profundidad. Ascidias, alcionarios, briozoos, poliquetos y otros han disminuido su presencia en la comunidad en más de un 50 %, mientras que la cobertura algal aumentó más de un 100 %

La información en el universo holográfico

Los resultados teóricos relativos a la entropía de los agujeros negros llevan a concluir que el universo podría ser un inmenso holograma

Jacob D. Bekenstein

Si preguntamos de qué se compone el mundo físico, se nos responderá que de "materia y energía". Pero quien sepa algo de ingeniería, biología y física nos citará también la información como elemento no menos importante. El robot de una fábrica de automóviles es de metal y plástico, pero no hará nada útil sin abundantes instrucciones que le digan qué pieza ha de soldar a otra. Un ribosoma de una célula se construye con aminoácidos y alimenta con la energía generada por la conversión del ATP en ADP, pero no podría sintetizar proteínas sin la información suministrada por el ADN del núcleo celular. Un siglo de investigaciones nos ha enseñado que la información desempeña una función esencial en los sistemas y procesos físicos. Hoy, una línea de pensamiento iniciada por John A. Wheeler, de la Universidad de Princeton, considera que el mundo físico está hecho de información; la energía y la materia serían accesorios.

Este punto de vista invita a reconsiderar cuestiones fundamentales. La capacidad de almacenamiento de la información de los discos duros y demás dispositivos de memoria ha ido creciendo a toda velocidad. ¿Cuándo se parará este progreso? ¿Cuál es la capacidad de información última de un dispositivo que pese, digamos, menos de un gramo y ocupe un centímetro cúbico (ése viene a ser el tamaño del chip de un ordenador)? ¿Cuánta información se necesita para describir todo un universo? ¿Podría tal descripción caber en la memoria de un ordenador? ¿Podríamos, tal como escribió William Blake, "ver el mundo en un grano de arena", o esas palabras sólo han de tomarse como una licencia poética?



Desarrollos recientes de la física teórica contestan algunas de estas preguntas; las respuestas podrían ser hitos importantes hacia la teoría definitiva de la realidad. Del estudio de las misteriosas propiedades de los agujeros negros se han deducido límites absolutos que acotan la información que cabe en una región del espacio o en una cantidad de materia y energía. Resultados ligados a ésos sugieren que nuestro universo, al que le percibimos tres dimensiones espaciales, podría en realidad estar "escrito" en una superficie bidimensional, como un holograma. Nuestra percepción ordinaria de un mundo tridimensional resultaría en tal caso una profunda ilusión o sólo una de dos maneras alternativas de ver la realidad. Quizás un grano de arena no abarque el mundo, pero sí lo haga una pantalla plana.

Las dos entropías

La teoría formal de la información nació de los artículos publicados en 1948 por el matemático estadounidense Claude E. Shannon. En ellos enunció la medida de la información más ampliamente usada hoy en día: la entropía. La entropía había venido siendo un concepto central de la termodinámica, la rama de la física que trata del calor. Suele decirse que la entropía termodinámica expresa el desorden de un sistema físico. En 1877 el físico austríaco Ludwig Boltzmann la caracterizó más precisamente como el número de estados microscópicos distintos en los que pueden hallarse las partículas que componen un trozo de materia de forma que siga pareciendo el mismo trozo desde un punto de vista macroscópico. En el caso del aire de una habitación, se contarían las maneras en que podrían distribuirse y moverse las moléculas de gas por la habitación.

Cuando Shannon buscó una manera de cuantificar la información contenida en un mensaje, la lógica le condujo a una fórmula que tenía el mismo aspecto que la de Boltzmann. La entropía de Shannon de un mensaje es el número de dígitos binarios, o bits, necesarios para codificarlo. No nos ilustra acerca del valor de la información, que depende mucho del contexto; pero en cuanto medida objetiva de la cantidad de información, la entropía de Shannon ha sido enormemente útil en ciencia y técnica. El diseño de todos los aparatos modernos de comunicación —desde los teléfonos portátiles hasta los módems y los reproductores de discos compactos— se basa en la entropía de Shannon.

La entropía termodinámica y la de Shannon son conceptualmente equivalentes: el número de configuraciones que se cuentan en la entropía de Boltzmann refleja la cantidad de in-

ALFRED T. KAMAJIAN

formación de Shannon que se necesitaría para realizar cualquier configuración determinada. Tales entropías presentan, sin embargo, dos diferencias principales. En primer lugar, la entropía termodinámica que emplean un químico o un experto en refrigeración se expresa en unidades de energía dividida por temperatura, mientras que la entropía de Shannon aplicada por un ingeniero de telecomunicaciones se da en bits, magnitud que carece de dimensiones. Esta diferencia no es más que una cuestión de convenciones.

Incluso cuando se las ha reducido a unidades comunes, los valores típicos de las dos entropías difieren mucho en magnitud. Un microchip de silicio que contenga un gigabyte de datos, por ejemplo, posee una entropía de Shannon de unos 10^{10} bits (un byte son ocho bits), muchísimo menor que la entropía termodinámica del chip, unos 10^{23} bits a temperatura ambiente. Esta discrepancia se debe a que esas entropías se calculan para grados de libertad diferentes. Un grado de libertad es cualquier cantidad que pueda cambiar, así una coordenada que especifica la localización de una partícula o una componente de su velocidad. La entropía de Shannon del chip sólo atiende al estado global de cada pequeño transistor impreso en el cristal de silicio: está *on* u *off*; representa un 0 o un 1 —un único grado de libertad binario—. La entropía termodinámica, por el contrario, depende de los estados de todos y cada uno de los miles de millones de átomos (con sus electrones en órbita) que forman cada transistor. A medida que la miniaturización nos acerque más al día en que cada átomo nos almacenará un bit de información, la entropía útil de Shannon del mejor microchip del momento se irá acercando a la entropía termodinámica de su material. Cuando las dos entropías se calculan para los mismos grados de libertad, resultan iguales.

¿Cuáles son los grados de libertad fundamentales? Al fin y al cabo, los átomos se componen de electrones y núcleos, los núcleos de protones y neutrones, y éstos de quarks. Muchos consideran hoy en día que los electrones y los quarks son excitaciones de supercuerdas, de las que piensan que son los entes fundamentales. Pero las vicisitudes de un siglo de revelaciones en la física nos han curado contra el dogmatismo. Podría haber más niveles de estructura en nuestro universo que los que sueña la física actual.

No se puede calcular la capacidad máxima de información de un pedazo de materia, o, de manera equivalente, su verdadera entropía termodinámica, sin saber la naturaleza de los últimos constituyentes de la materia, de su nivel de estructura más profundo, al que llamaré "nivel X". (Esta ambigüedad no le causa problemas a la termodinámica práctica, la de un motor de coche, por ejemplo, ya que puede ignorar los quarks del interior del átomo; no cambian de estado en las moderadas condiciones del motor.) Dado el vertiginoso progreso de la miniaturización, juguemos a imaginar un día en que los quarks sirviesen para almacenar información, quizás un bit cada uno. ¿Cuánta información cabría entonces en nuestro cubo de un centímetro de lado? ¿Y cuánta si lográsemos controlar las supercuerdas, o niveles más profundos aún ni soñados? Sorprendentemente, los desarrollos en la física de la gravitación en los treinta últimos años han proporcionado algunas respuestas claras a preguntas que parecían tan inabordables.

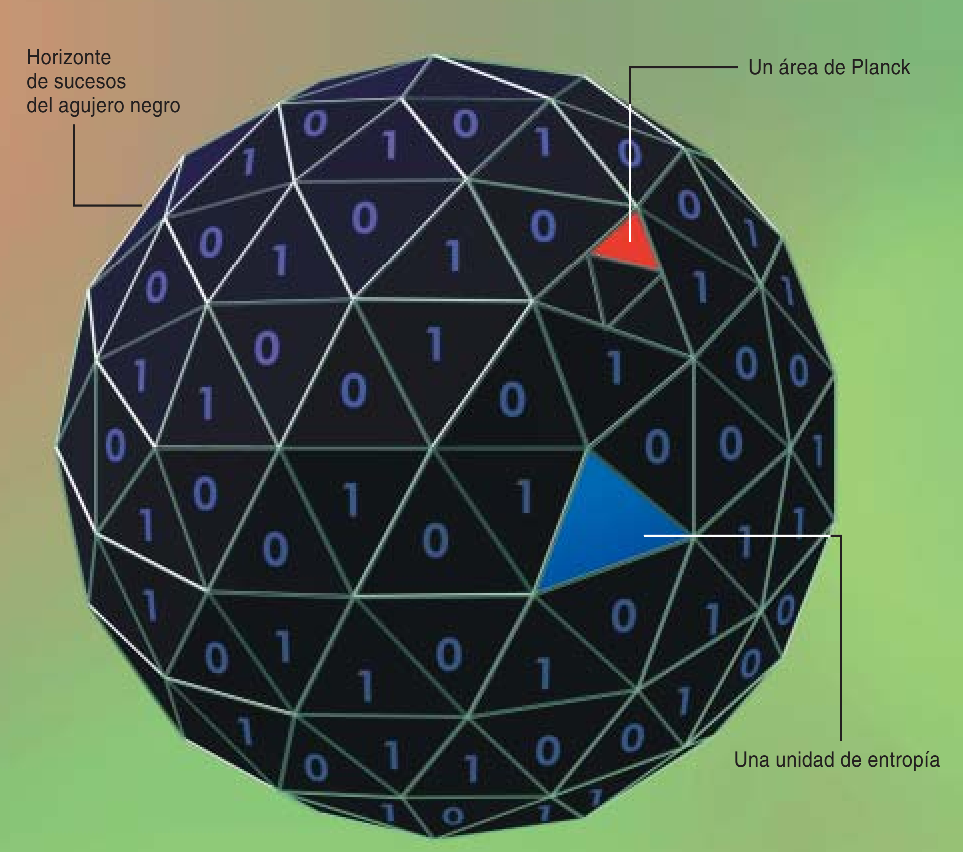
Resumen/*El mundo como holograma*

- Una hipótesis asombrosa, el principio holográfico, mantiene que el universo es como un holograma: igual que jugando con luces se registra una imagen tridimensional en un pedazo de película plana, este universo que nos parece tridimensional podría equivaler por completo a otro conjunto de campos cuánticos y leyes físicas "pintado" en una superficie vasta y lejana.
- La física de los agujeros negros —concentraciones de masa con una densidad inmensa— da a entender la veracidad del principio. El estudio de los agujeros negros enseña, por mucho que desafíe al sentido común, que el máximo contenido de información de cualquier región del espacio viene dado, no por su volumen, sino por el área de la superficie que la delimita.
- Los físicos esperan que este sorprendente hallazgo sea una clave en la búsqueda de la definitiva teoría de la realidad.

Termodinámica del agujero negro

Protagonista de estos avances es el agujero negro. Los agujeros negros son una consecuencia de la relatividad general, la teoría geométrica de la gravitación establecida por Albert Einstein en 1915. Según esta teoría, la gravitación surge de la curvatura del espacio-tiempo, que hace que los objetos se muevan como si estuviesen atraídos por una fuerza. A la inversa, la causa de la curvatura es la presencia de materia y energía. Según las ecuaciones de Einstein, una concentración suficientemente densa de materia o energía curva tanto el espacio-tiempo, que lo rasga y nace un agujero negro. Las leyes de la relatividad prohíben a todo lo que caiga en un agujero negro volver a salir; por lo menos, dentro de la formulación clásica (es decir, no cuántica) de la física. El punto de no retorno, el horizonte de sucesos del agujero negro, es de crucial importancia. En el caso más simple, se trata de una esfera; tanta más área tendrá cuanto mayor sea la masa del agujero.

Es imposible determinar lo que hay dentro de un agujero negro. Ninguna información detallada puede emerger del horizonte y escapar al mundo exterior. Con todo, un pedazo de materia deja algunos rastros cuando desaparece para siempre en un agujero negro. Su energía (contamos cualquier masa como energía, de acuerdo con la fórmula de Einstein $E = mc^2$) queda permanentemente reflejada en un incremento de la masa del agujero negro. Si la materia que absorbe orbitaba a su alrededor, el momento angular correspondiente se añadirá al momento angular del agujero negro. Tanto la masa como el momento angular del agujero negro son mensurables gracias a sus efectos en el



1. LA ENTROPIA DE UN AGUJERO NEGRO es proporcional al área de su horizonte de sucesos (superficie de cuyo interior no puede escapar, a causa de la gravedad del agujero, ni siquiera la luz). En concreto, un agujero con un horizonte que abarque A áreas de Planck tendrá A/4 unidades de entropía. (El área de Planck, unos 10^{-66} centímetros cuadrados, es la unidad cuántica de área fundamental, determinada por la intensidad de la gravedad, la velocidad de la luz y el tamaño de los cuantos.) Como si la entropía, en cuanto medida de la información, estuviese escrita sobre el horizonte de sucesos, de suerte tal que cada bit (cada 0 o 1 de la codificación digital) correspondiera a 4 áreas de Planck.

ALFRED T. KAMAJIAN

espacio-tiempo de los alrededores del agujero; así, los agujeros negros respetan las leyes de conservación de la energía y del momento angular. En cambio, parece como si otra ley fundamental, la segunda ley de la termodinámica, se violase.

La segunda ley de la termodinámica compendia algo conocido por todos: que la mayoría de los procesos naturales son irreversibles. Una taza de té cae de la mesa y se rompe; nadie ha visto jamás que los trozos salten del suelo y recompongan la taza. La segunda ley de la termodinámica prohíbe la inversión del proceso. Establece que la entropía de un sistema físico aislado nunca decrece; en el mejor de los casos, permanecerá constante; por lo normal, aumentará. Esta ley es esencial para la físico-química y la ingeniería; cabe sostener que es la ley física que más se tiene en cuenta fuera de la propia física.

Tal y como explicó Wheeler, con la materia que cae en un agujero negro desaparece también su entropía; diríase que así se supera y arrumba la segunda ley. Una idea de cómo podía resolverse este problema llegó en 1970. Demetrious Christodoulou, entonces estudiante de doctorado de Wheeler en Princeton, y por otra parte Stephen W. Hawking, de la Universidad de Cambridge, demostraron que en varios procesos, entre ellos la fusión de dos agujeros negros, nunca decrecía el área total de los horizontes de sucesos. La analogía con la tendencia de la entropía a aumentar me llevó a proponer en 1972 que un agujero negro tiene una entropía proporcional al área de su horizonte (*véase la figura 1*). Conjeturé que, cuando la materia cae en un agujero negro, el aumento de la entropía de éste siempre compensa, con creces incluso, la entropía "perdida" por la materia. Con más generalidad: la suma de la entropía del agujero negro y de la entropía ordinaria fuera del mismo no puede decrecer. Esta es la generalización de la segunda ley (o GSL).

La GSL ha superado un gran número de estrictas pruebas, si bien puramente teóricas. Cuando una estrella se desploma sobre sí misma y crea un agujero negro, la entropía de éste supera en mucho la de la estrella. En 1974 Hawking demostró que un agujero negro emite espontáneamente radiación térmica —hoy denominada "radiación de Hawking"— mediante un proceso cuántico. El teorema de Christodoulou-Hawking falla ante ese fenómeno (la masa del agujero negro y, por tanto, el área de su horizonte decrecen), pero la GSL ataja el problema: la entropía de la radiación emergente compensa la merma de la entropía del agujero negro, de manera que se conserva la GSL. En 1986 Rafael D. Sorkin, de la Universidad de Syracuse, se apoyó en que el horizonte bloquee la información interna del agujero e impida que influya en el exterior para demostrar que la GSL (o algo muy parecido a ella) tiene que ser válida en cualquier proceso concebible que sufran los agujeros negros. Su profundo argumento dejaba claro que la entropía a que se refiere la GSL coincide con la calculada en el nivel X, sea cual sea ese nivel.

Con su proceso de radiación Hawking determinó la constante de proporcionalidad entre la entropía de un agujero negro y el área del horizonte: la entropía del agujero negro es exactamente una cuarta parte del área del horizonte de sucesos medida en áreas de Planck. (La longitud de Planck, unos 10^{-33} centímetros, es la escala de longitud fundamental relacionada con la gravedad y la mecánica cuántica. Su cuadrado es el área de Planck.) Incluso desde un punto de vista termodinámico, se trata de una enorme cantidad de entropía. La entropía de un agujero negro de un centímetro de diámetro sería de unos 10^{66} bits, aproximadamente igual a la entropía termodinámica de un cubo de agua de 10.000 millones de kilómetros de lado.

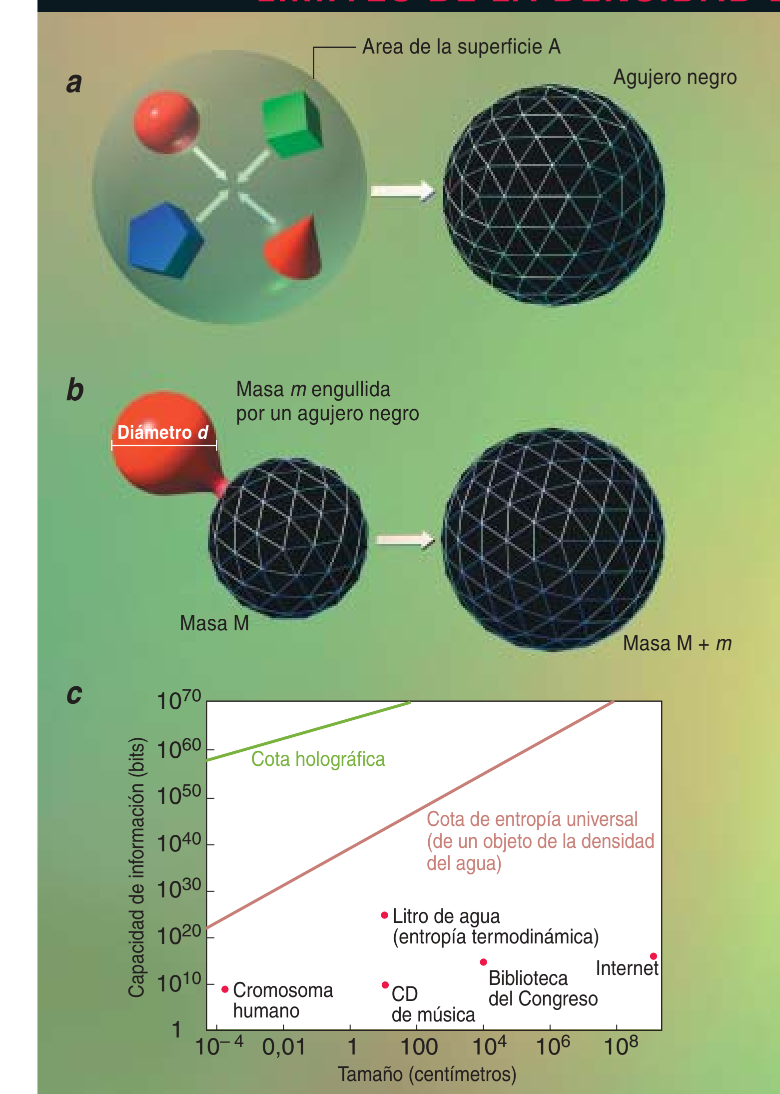
El mundo, un holograma

La GSL nos permite establecer cotas de la capacidad de información de cualquier sistema físico aislado, límites que se refieren a la información en todos los niveles de estructura, hasta el X. En 1980 empecé a estudiar la primera de tales cotas la "cota de entropía universal", así se llama, limita la entropía que puede contener una determinada masa de un tamaño concreto. Una idea afín, la cota holográfica, fue apuntada en 1995 por Leonard Susskind, de la Universidad de Stanford. Limita cuánta entropía pueden contener la materia y la energía que ocupen un determinado volumen de espacio.

En su trabajo sobre la cota holográfica, Susskind consideró cualquier masa aislada, aproximadamente esférica, que no fuese un agujero negro y que cupiera dentro de una superficie cerrada de área *A*. Si la masa puede contraerse hasta convertirse en un agujero negro, éste acabará con un horizonte de área menor que *A*; su entropía será, por tanto, menor que A/4. Según la GSL, la entropía del sistema no puede decrecer; por tanto, la entropía de la masa original no pudo haber sido mayor que A/4. En consecuencia, la entropía de un sistema físico aislado limitado por un área *A* es necesariamente menor que A/4. ¿Qué sucede si la masa no sufre un colapso gravitatorio espontáneo que la convierta en un agujero negro? En el año 2000 demostré que con un pequeño agujero negro podía convertirse el sistema en un agujero negro no muy diferente del considerado en el argumento de Susskind. La cota es, por tanto, independiente de la constitución del sistema o de la naturaleza del nivel X. Sólo depende de la GSL.

Podemos ahora responder alguna de las cuestiones espinosas sobre los últimos límites del almacenamiento de información. Un dispositivo que mida un centímetro de diámetro podría, en principio, contener hasta 10^{66} bits. El universo visible contiene por lo menos 10^{100} bits de entropía, que cabría en principio almacenar dentro de una esfera de una décima de año luz de diámetro. Estimar la entropía del

LIMITES DE LA DENSIDAD DE INFORMACION



La termodinámica de los agujeros negros permite deducir cotas de la densidad de entropía o información en diversas circunstancias.

La cota holográfica define cuánta información puede albergar una determinada región del espacio. Para averiguarlo, se considera una distribución de materia aproximadamente esférica incluida en el interior de una superficie de área A. Se provoca que esa materia colapse y forme un agujero negro (*a*). El área del agujero negro será menor que A; su entropía no deberá, pues, llegar a A/4. Debido a que la entropía no puede menguar, se infiere que la distribución original de materia también deberá tener menos de A/4 unidades de entropía o información. Este resultado —que el máximo contenido de información de una cierta región del espacio viene determinado por su área— desafía la expectativa del sentido común, a tenor del cual la capacidad de una región debería depender de su volumen.

La cota universal de la entropía define cuánta información puede contener una masa *m* de diámetro *d*. Para deducirla, hay que imaginar que una cápsula de materia es engullida por un agujero negro no mucho mayor que ella (*b*). El aumento del tamaño del agujero negro establece un límite a cuánta entropía pudo haber contenido la cápsula. Este límite es más estricto que la cota holográfica, salvo cuando la densidad de la cápsula se aproxima a la del agujero negro, en cuyo caso las dos cotas coinciden.

La capacidad de almacenamiento de datos de cualquier técnica existente se halla muy lejos de ambas cotas, que superan también en mucho a la densidad de información de los cromosomas y la entropía termodinámica del agua (*c*).

—J.D.B.

LAURIE GRACE (*gráfica*)


2. EL CONTENIDO DE INFORMACION de un montón de chips de ordenador aumenta en proporción al número de chips; dicho de otro modo: aumenta con el volumen que ocupe. Esta simple regla fallará para un montón de cierto tamaño: en cierto momento, la información superará la cota holográfica, que la liga al área del montón, no al volumen. El "fallo" se producirá cuando el montón de chips, inmenso, se derrumbe sobre sí mismo y forme un agujero negro.

ALFRED T. KAMAJIAN

universo es un problema difícil, sin embargo; parecen verosímiles números mucho mayores, que requerirían una esfera casi tan grande como el propio universo.

Pero hay otro aspecto de la cota holográfica que sorprende: la máxima entropía posible depende del área de la superficie que delimita el volumen y no de éste. Imaginemos que apilamos chips de memoria de ordenador. El número de transistores —la capacidad total de almacenamiento de datos— aumenta con el volumen del montón. Con éste, aumentará también la entropía termodinámica total de los chips. Pero es bien notable que la teórica capacidad final de información del espacio ocupado por el montón sólo aumente con el área superficial. Como el volumen aumenta más deprisa que el área, llegará un momento en que la entropía de los chips sobrepasará la cota holográfica. Parecería entonces que, o bien la GSL, o bien nuestras ideas corrientes de entropía y de capacidad de información, fundadas en el sentido común, tendrían que fallar. Pero lo que fallaría sería el propio montón: se derrumbaría bajo su propia gravedad y formaría un agujero negro antes de que se alcanzara el callejón sin salida. Cada chip adicional de memoria aumentaría la masa y el área superficial del agujero negro de manera que se respetase la GSL.

Este sorprendente resultado —que la capacidad de información depende del área superficial— tiene una explicación natural si es cierto el *principio* holográfico (propuesto en 1993 por el Nobel Gerard 't Hooft, de la Universidad de Utrecht, y elaborado por Susskind). En el mundo ordinario, un holograma es una clase especial de fotografía que genera una auténtica imagen tridimensional cuando se la ilumina de la manera adecuada. Toda la información que describe la imagen tridimensional va codificada en un diagrama de luz y zonas oscuras impreso en un pedazo de película bidimensional, lista para que se la regenere. El principio holográfico afirma que un análogo de esta maravilla visual es aplicable a la plena descripción física de cualquier sistema que ocupe una región tridimensional: otra teoría física, definida sólo en el límite bidimensional de la región, describirá por completo la física tridimensional. Si un sistema tridimensional completo puede ser plenamente descrito por una teoría física definida sólo en su contorno bidimensional, se esperará que el contenido de información del sistema no exceda del contenido de la descripción limitada al contorno.

El autor

JACOB D. BEKENSTEIN ha contribuido a la fundación de la termodinámica de los agujeros negros y a elucidar otros aspectos de la conexión entre información y gravedad. Desempeña la cátedra Polak de física teórica de la Universidad Hebrea de Jerusalén. Dedica este artículo a John Archibald Wheeler, que dirigió su tesis hace 30 años.

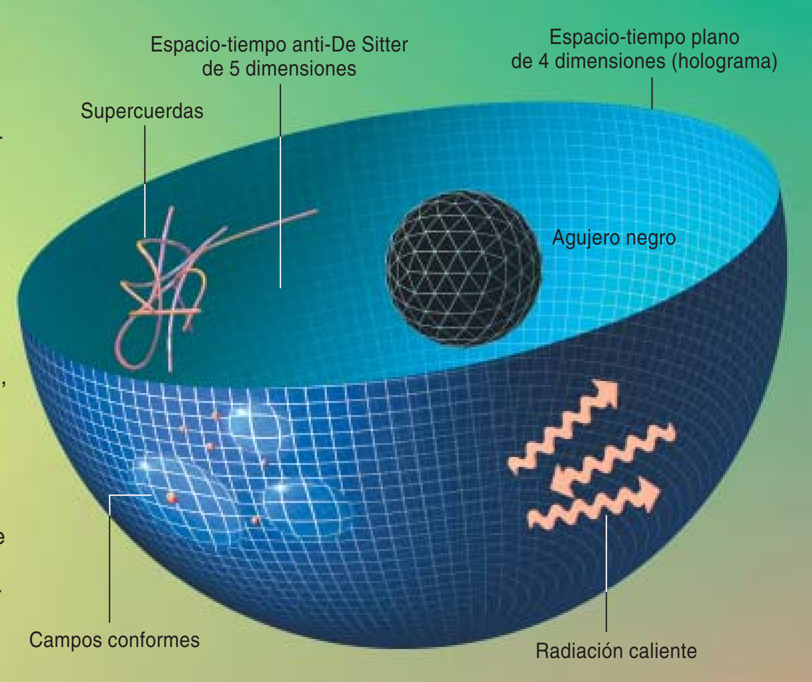
Un universo pintado sobre su contorno

¿Podemos aplicar el principio holográfico a todo el universo? El universo real es un sistema tetradimensional: tiene volumen y se extiende en el tiempo. Si la física de nuestro universo fuese holográfica, habría otro conjunto de leyes físicas, que se aplicaría en algún contorno tridimensional del espacio-tiempo, equivalente a nuestra conocida física tetradimensional. Aún no sabemos de ninguna tal teoría tridimensional que actúe de esta manera. Y ¿qué superficie deberíamos usar como contorno del universo? Estudiar modelos más sencillos que nuestro universo real es un paso hacia la realización de estas ideas.

Una clase de ejemplos concretos del principio holográfico en acción lo constituyen los espacio-tiempos anti-De Sitter. El espacio-tiempo de De Sitter original consiste en un modelo de universo obtenido por Willem de Sitter en 1917, una solución de las ecuaciones de Einstein que incluye una fuerza repulsiva, la "constante cosmológica". El espacio-tiempo de De Sitter está vacío, se expande a un ritmo acelerado y es muy simétrico. En 1997, de la observación de explosiones de supernova distantes se dedujo que nues-

UN ESPACIO-TIEMPO HOLOGRAFICO

El principio holográfico convierte en equivalentes a dos universos de distinta dimensión que obedecen leyes físicas dispares. Se ha encontrado matemáticamente un ejemplo de esta situación en cierto tipo de espacio-tiempo (el anti-De Sitter) de cinco dimensiones y su contorno de cuatro dimensiones. El universo pentadimensional queda registrado como un holograma en su superficie periférica tetradimensional. La teoría de supercuerdas rige el espacio-tiempo pentadimensional y una "teoría de campos conformes con partículas puntuales", el holograma tetradimensional. Un agujero negro del espacio-tiempo pentadimensional equivale a radiación caliente en el holograma; de ese modo, el agujero y la radiación tienen la misma entropía a pesar de que el origen físico de ésta difiere por completo en cada caso. Aunque estas dos descripciones del universo parecen distintas, ningún experimento podría diferenciarlas, ni siquiera en principio.



ALFRED T. KAMAJIAN

tro universo se expande aceleradamente; es probable que en el futuro se vaya pareciendo cada vez más a un espacio-tiempo de De Sitter. Pero si en las ecuaciones de Einstein la repulsión se convierte en atracción, de la solución de De Sitter se pasa a un espacio-tiempo anti-De Sitter, no menos simétrico y que cuenta con un contorno localizado en "el infinito" que se parece mucho a nuestro espacio-tiempo habitual.

Con el espacio-tiempo anti-De Sitter, los teóricos han encontrado una realización concreta del principio holográfico: un universo descrito por la aplicación de la teoría de supercuerdas a un espacio-tiempo anti-De Sitter equivale a una teoría cuántica de campos que opera en la frontera de dicho espacio-tiempo. Así, toda la majestad de la teoría de supercuerdas en un universo anti-De Sitter está pintada sobre el contorno del universo. Juan Maldacena, ahora en la Universidad de Harvard, conjeturó tal relación en 1997 para un anti-De Sitter pentadimensional; luego lo confirmaron para muchas otras situaciones Edward Witten, del Instituto de Estudios Avanzados de Princeton, y Steven S. Gubser, Igor R. Klebanov y Alexander M. Polyakov, de la Universidad de Princeton. Ahora se conocen ejemplos de esta correspondencia holográfica para espacio-tiempos de diversas dimensiones.

Se establece así la equivalencia de dos teorías, en apariencia muy diferentes, que ni siquiera actúan en espacios de la misma dimensión. Las criaturas que vivieran en uno de estos universos serían incapaces de determinar si habitaban en un universo pentadimensional descrito por una teoría de cuerdas o en uno tetradimensional descrito por una teoría cuántica de campos con partículas puntuales. (Las estructuras de sus cerebros les podrían inducir un irresistible prejuicio de "sentido común" a favor de una u otra descripción, de la misma manera que nuestros cerebros construyen una innata percepción de que nuestro universo tiene tres dimensiones espaciales.)

Gracias a la equivalencia holográfica, en vez de un difícil cálculo en el espacio-tiempo frontera tetradimensional, como los relativos a quarks y gluones, puede efectuarse otro más sencillo en el muy simétrico espacio-tiempo anti-De Sitter de cinco dimensiones. La correspondencia funciona también en el otro sentido. Witten ha demostrado que un agujero negro en un espacio-tiempo anti-De Sitter corresponde a una radiación caliente en la física alternativa que opera en el espacio-tiempo frontera. La entropía del agujero, un concepto profundamente misterioso, es igual a la entropía de la radiación, que es muy corriente.

El universo en expansión

Altamente simétrico y vacío, el universo anti-De Sitter pentadimensional es muy distinto de nuestro universo tetradimensional, lleno de materia y radiación, escenario de sucesos violentos. Incluso cuando describimos de manera sólo aproximada nuestro universo real como uno dotado de materia y radiación distribuidas uniformemente, no obtenemos un universo anti-De Sitter, sino un universo de Friedmann-Robertson-Walker (FRW). Hoy en día, la mayoría de los cosmólogos coinciden en que nuestro universo se parece a un FRW, mundo infinito, sin frontera, que se irá expandiendo por siempre jamás.

¿Concuerda un tal universo con el principio holográfico o con la cota holográfica? El argumento de Suss-

kind, basado en un colapso creador de un agujero negro, no nos ayuda aquí. La cota holográfica deducida de los agujeros negros debe fallar en un universo uniforme en expansión. La entropía de una región uniformemente llena de materia y radiación es, sin duda, proporcional a su volumen. Una región lo bastante grande violaría por tanto la cota holográfica.

En 1999 Raphael Bousso, entonces en Stanford, propuso una cota holográfica modificada; se ha ido comprobando su validez incluso en situaciones en las que no pueden aplicarse las cotas de las que hemos hablado antes. La formulación de Bousso parte de cualquier superficie bidimensional adecuada, cerrada como una esfera o abierta como una hoja de papel. Hay que imaginar a continuación breves destellos de luz que salgan simultánea y perpendicularmente de toda una cara de la superficie. La única condición es que esos imaginarios rayos de luz converjan. La luz emitida por la superficie interior de una capa esférica, por ejemplo, satisface este requisito. Se considera entonces la entropía de la materia y la radiación que atraviesan estos imaginarios rayos hasta llegar a los puntos donde se cruzan. Bousso conjeturó que no puede superar la representada por la superficie inicial: un cuarto de su área, medida en áreas de Planck. Esta forma de medir la entropía difiere de la usada en la cota holográfica original. La cota de Bousso no se refiere a la entropía de una región en un tiempo, sino a la suma de entropías de lugares en una diversidad de tiempos: los que son "iluminados" por los destellos de luz que salen de la superficie.

La cota de Bousso subsume otras cotas de entropía, a la vez que evita sus limitaciones. Tanto la cota de entropía universal como la forma de 't Hooft-Susskind de la cota holográfica se pueden deducir de la de Bousso para cualquier sistema aislado que no evolucione rápidamente y cuyo campo gravitacional no sea intenso. Cuando estas condiciones se superan —como en el caso de una esfera de materia que colapsa ya dentro de un agujero negro—, aquellas cotas fallan, mientras que la de Bousso se mantiene.



3. NUESTRA PERCEPCION INNATA de que el mundo es tridimensional podría resultar una extraordinaria ilusión.

ALFRED T. KAMAJIAN

Bousso también ha demostrado que su estrategia sirve para encontrar superficies bidimensionales en las que se puedan inscribir los hologramas del mundo.

Augurios de una revolución

Se han propuesto muchas otras cotas de la entropía. La proliferación de variaciones de la idea holográfica deja claro que no ha alcanzado aún la categoría de ley física. Pero aunque no se entienda todavía del todo el planteamiento holográfico, parece que arraigará. Y de su mano viene la convicción de que deberá abandonarse la creencia fundamental, imperante durante medio siglo, en que la teoría de campos es el lenguaje final de la física. Los campos, el electromagnético por ejemplo, varían continuamente de un punto a otro; comprenden por tanto una infinidad de grados de libertad. La teoría de supercuerdas también abarca un número infinito de grados de libertad. La holografía restringe los grados de libertad del interior de una superficie delimitadora a un número finito; la teoría de campos, con su infinitud, no puede ser la última palabra. Aun cuando se controlase la infinitud, habría que encajar la misteriosa ligazón entre la información y el área.

La holografía puede ser una guía hacia una teoría mejor. ¿De qué tipo será la teoría fundamental? La cadena de razonamientos relativa a la holografía sugiere a algunos, especialmente a Lee Smolin, del Instituto Perímetro de Física Teórica, en Waterloo, que no deberá referirse a campos, ni tampoco al espaciotiempo, sino al intercambio de información entre procesos físicos. Si es así, la idea de que el mundo está hecho de información habrá encontrado un digno vehículo.

Bibliografía complementaria

BLACK HOLE THERMODYNAMICS. Jacob D. Bekenstein en *Physics Today*, vol. 33, n.º 1, pág. 24-31; enero 1980.

BLACK HOLES AND TIME WARPS: EINSTEIN'S OUTRAGEOUS LEGACY. Kip S. Thorne. W. W. Norton, 1995.

LOS AGUJEROS NEGROS Y LA PARADOJA DE LA INFORMACIÓN. Leonard Susskind, en *Investigación y Ciencia*, págs. 12-18; junio 1997.

THE UNIVERSE IN A NUTSHELL. Stephen Hawking. Bantam Books, 2001.

THREE ROADS TO QUANTUM GRAVITY. Lee Smolin. Basic Books, 2002.



El oráculo de Delfos

Los griegos de la Antigüedad clásica andaban en lo cierto cuando afirmaban que los vapores procedentes del interior de la Tierra inspiraban a las sacerdotisas

John R. Hale, Jelle Zeilinga De Boer, Jeffrey P. Chanton y Henry A. Spiller

1. LA PITONISA DE DELFOS aparece en esta interpretación fotográfica inhalando vapores. Recientes investigaciones corroboran lo que decían los antiguos: la inspiración oracular provenía de unos gases que intoxicaban a la médium. En la realidad, los gases habrían sido invisibles.

SANJAY KOTHARI; TRUCOLLAGE

El templo de Apolo, enclavado en el espectacular paraje montañoso de Delfos, fue el más importante centro religioso del antiguo mundo griego porque allí estaba el poderoso oráculo. Los generales le consultaban sobre estrategia; los colonizadores le pedían consejo antes de partir hacia Italia, España y Africa; los particulares le preguntaban acerca de problemas de salud y de negocios. Las respuestas del oráculo ocupan un lugar muy destacado en los mitos: cuando Orestes preguntó si debía vengarse de su madre por haber asesinado ésta a su padre, el oráculo le animó a ello. Edipo, habiéndole el oráculo avisado de que mataría a su padre y se casaría con su madre, procuró, con el poco éxito que sabemos, evitar su sino.

El oráculo de Delfos actuaba en un lugar concreto, el *adyton*, un lugar reservado del interior del templo vedado a los profanos; intervenía, por mediación de un ser humano; éste, a la manera de una medium poseída, hablaba por Apolo, el dios de la profecía. Cosa singular en la misógina Grecia, ese portavoz era una mujer, la *pythia* o pitonisa. Y, a diferencia de la mayoría de los sacerdotes y sacerdotisas griegos, no heredaba su oficio por su parentesco con familias de la nobleza. Aunque tenía que ser natural de Delfos, podía ser vieja o joven, rica o pobre, bien instruida o iletrada. Debía pasar un largo período de intensa preparación, ayudada por una hermandad de mujeres délficas que mantenían en el templo el eterno fuego sagrado.

La explicación clásica

La tradición atribuía la inspiración profética del poderoso oráculo a fenómenos geológicos: una profunda grieta del terreno, un vapor que salía de ella y una fuente. Hará alrededor de un siglo, los estudiosos rechazaron esta explicación porque excavando allí los arqueólogos no encontraron ni grieta ni vapor alguno. Sin embargo, el antiguo testimonio, muy difundido, procedía de diversas fuentes: de historiadores como Plinio y Diodoro, de

filósofos como Platón, de poetas como Esquilo, de oradores como Cicerón, del geógrafo Estrabón, del escritor viajero Pausanias e incluso de un sacerdote de Apolo que ofició en Delfos, el famoso ensayista y biógrafo Plutarco.

Estrabón (64 a.C.-25 d.C.) escribió: "Dicen que la sede del oráculo es una caverna profundamente hundida en el terreno y por cuya boca, más bien estrecha, sale un *pneuma* (gas, vapor, exhalación; de ahí nuestras palabras 'neumático' y 'neumonía') que produce el trance o la posesión divina. Encima de esta estrecha abertura hay instalado un trípode; sentándose en él, la pitonisa inhala el vapor y profetiza".

Plutarco (46-120 d.C.) describió por extenso el funcionamiento del oráculo, del que fue testigo ocular. Explicó las relaciones entre el dios, la mujer y el gas comparando a Apolo con un músico, a la mujer con el instrumento y al *pneuma* con el plectro con que el dios la tocaba para que hablara. Pero Plutarco insiste en que el *pneuma* constituía sólo un detonante; si la mujer elegida podía responder cuando se exponía al *pneuma* era gracias a la preparación y purificación previas (que sin duda incluían la abstinencia sexual y posiblemente el ayuno). Una persona corriente olería el gas, pero no por ello entraría en trance oracular.

Plutarco dio también algunas características físicas del *pneuma*: despedía el olor de un perfume suave y surgía "como de una fuente" en el *adyton* donde se sentaba la pitia, pero en ocasiones los sacerdotes y los consultantes podían olerlo en la antecámara donde esperaban las respuestas. Manaba como un gas libre o en el agua. En la época de Plutarco, la emisión del aroma se había vuelto débil e irregular; ésa era la causa, en su opinión, de que se estuviese debilitando la influencia del oráculo en los asuntos del mundo. Conjeturaba Plutarco que quizá la esencia vital se iba agotando, que fuertes aguaceros la habían diluido o que un gran terremoto ocurrido más de cuatro siglos atrás había cegado en parte la abertura por donde salía. Pudiera ser —seguía conjeturando— que el vapor hubiese encontrado una nueva salida. Estas hipótesis sobre la merma de la emisión evidencian que Plutarco creía que el gas se originaba en la roca sobre la que se asentaba el templo.

Una generación más tarde el viajero Pausanias repite a Plutarco al mencionar que el *pneuma* surge del agua. Refiere que vio en la ladera, más arriba del templo, el manantial Kassotis, del que había oído que se hundía en la tierra y volvía a emerger en el *adyton*, donde sus aguas hacían que las mujeres profetizaran.

Plutarco y otras fuentes indican que, durante las sesiones normales, la pitonisa entraba en un leve trance; podía mantenerse derecha sentada en el trípode y pasarse allí bastante rato (aunque tal vez tuvieran que relevarla una segunda pitia y hasta una tercera cuando era larga la fila de consultantes); escuchaba las pre-

2. LA UNICA REPRESENTACION QUE SE HA CONSERVADO de la sacerdotisa, o pitonisa, de Delfos creada en la misma época del oráculo nos muestra la cámara de baja techumbre y a la pitonisa sentada en un trípode. La mujer tiene en una mano una ramita de laurel (árbol sagrado de Apolo) y, en la otra, una vasija que seguramente contenía agua de una fuente que burbujeaba en la cámara y portaba los gases que inducían el trance. En esta escena mitológica, el rey Egeo de Atenas consulta a la primera pitonisa, Temis. Cerámica decorada por un alfarero ateniense hacia el 440 a.C.

BRIDGEMAN ART LIBRARY

Resumen/*Intoxicación y profecía*

- Durante el siglo pasado los estudiosos desacreditaron como mero mito la explicación tradicional: que en Delfos salían de una grieta del suelo unos vapores que intoxicaban e inspiraban a las sacerdotisas.
- La investigación reciente muestra que aquella explicación era extraordinariamente certera.
- Los autores del presente artículo han identificado dos fallas geológicas que se cruzan bajo el emplazamiento del oráculo.
- Más aún, es muy probable que de los estratos calizos de la zona, ricos en elementos petroquímicos, emanara etileno, gas que provoca estados que recuerdan a un trance y podría haber surgido a través de las grietas abiertas por las fallas.

guntas y daba respuestas inteligibles. Durante las sesiones oraculares, la pitonisa hablaba con una voz alterada y tendía a canturrear sus respuestas; se permitía retruécanos e ingeniosidades. Al final, según Plutarco, quedaba la pitonisa en un estado parecido al del corredor después de la carrera o al de quien ha participado en una danza extática.

En una ocasión en la que o Plutarco mismo, o uno sus colegas se hallaron presentes, las autoridades del templo, para quedar bien con los miembros de una importante embajada, forzaron a la pitonisa a profetizar aunque el día no era propicio. Bajó contra su voluntad al subterráneo *adyton* e inmediatamente fue arrebatada por un poderoso espíritu maligno. Estando posesa, en vez de hablar o entonar según lo acostumbrado, gruñó y lanzó alaridos, se agitó con enorme violencia y se abalanzó contra las puertas, donde se desmayó. Aterrorizados, los sacerdotes y los consultantes escaparon de allí, pero al cabo de un rato volvieron y la levantaron. A los pocos días la mujer murió.

La nueva tradición

Generaciones de eruditos tuvieron por válidos estos testimonios hasta que un joven clasicista inglés, Adolphe Paul Oppé, visitó, allá por 1900, las excavaciones que dirigían en Delfos unos arqueólogos franceses. No logró ver sima alguna, ni oír nada acerca de emanaciones de gases, y publicó un influyente artículo, con tres tajantes aseveraciones: la primera, que en el templo de Delfos jamás había habido ni simas, ni emisiones de gas; la segunda, que, aun cuando las hubiese habido, ningún gas natural producía un estado parecido al trance espiritista; y la tercera, que lo que Plutarco contaba de una pitonisa que sufrió un violento frenesí y murió poco después no concordaba con la sólita descripción de la pitonisa sentada en el trípode y canturreando sus vaticinios. Oppé concluía que se podía descartar todo el testimonio antiguo.

Con su crítica dejó Oppé inerme al mundo académico. Sus opiniones estaban expresadas con tanta contundencia, que pasaron a ser la nueva ortodoxia. El no haberse hallado la gran abertura que los arqueólogos esperaron encontrar parecía dar respaldo a sus argumentos. Su teoría recibió un apoyo más cuando, en 1950, el arqueólogo francés Pierre Amandry añadió que sólo una zona volcánica —y Delfos no lo era— podría haber producido un gas como el descrito en las fuentes clásicas. El caso parecía cerrado. La tradición original de los autores griegos y latinos siguió viva tan sólo en los libros de divulgación y en las explicaciones de los guías locales, cuyos predecesores, según Oppé, dieron en su día origen al mito de la sima y el vapor.

La situación cambió en el decenio de 1980, cuando un proyecto de desarrollo patrocinado por la ONU emprendió en Grecia un estudio de las fallas activas (aquellas a lo largo de las cuales se han producido terremotos en los últimos siglos). Uno de nosotros (De Boer, que es geólogo) participó en el proyecto; observó que había dos caras de falla expuestas al este y al oeste del santuario. Las interpretó como señales de una línea de falla que se extendía a lo largo de la ladera sur del monte Parnaso y bajo el emplazamiento del oráculo. Pero, conocedor de la tradición clásica e ignorante de que el moderno escepticismo la desechaba, no dio especial importancia a su observación.

Más de una década después conocería De Boer a otro miembro de nuestro equipo (Hale) en un yacimiento arqueológico de Portugal. Hale, que es arqueólogo, preguntó a De Boer qué opinaba como geólogo de los indicios de daños causados por un terremoto en una antigua villa romana. De Boer mencionó que había visto la falla que corría por debajo del templo de Delfos. Conscientes de la importancia que aquella observación tenía para la interpretación acertada de los antiguos relatos, decidieron formar un equipo que explorase a fondo el lugar.



CENTRO MEDICO PRESBITERIANO SAN LUCAS-UNIVERSIDAD RUSH

3. LOS EXPERIMENTOS con anestesia dirigidos por Isabella Herb (*de pie*) a mediados del siglo XX fueron un elemento de juicio fundamental en la deducción de qué gas pudo emitirse bajo el templo de Delfos. Herb y sus colaboradores descubrieron que las bajas concentraciones de etileno provocan un estado parecido al de un trance.

Vuelta a la explicación clásica

Durante nuestro primer estudio sobre el terreno, en 1996, llevamos a cabo prospecciones geológicas y examinamos los cimientos del templo sacados a la luz por los arqueólogos franceses. Tiene el templo unas cuantas características anómalas que, aun en el supuesto de que no se hubiesen conservado los relatos de Plutarco y demás autores, seguirían requiriendo alguna interpretación especial de su función. La primera de esas características es que el interior del santuario esté hundido de dos a cuatro metros por debajo del nivel del suelo cir-

cundante. La segunda, su asimetría: una interrupción de la columnata interna parece dejar sitio para alguna construcción o elemento ya desaparecido. La tercera, que entre los cimientos, cerca de la parte reservada, se construyó una complicada canalización para agua de manantial, junto a otros pasajes subterráneos. Así, el templo de Apolo parecía diseñado para cerrar un terreno donde manaba una fuente, más que para la función normal de un templo: albergar la imagen de un dios.

Durante aquella primera exploración cartografiamos la principal línea de falla, la falla de Delfos; era la que De Boer había observado durante su anterior inspección y que atravesaba el santuario de este a oeste. Algo después descubriríamos la cara expuesta de una segunda falla, en una quebrada que había más arriba del templo. Esta segunda línea, a la que llamamos falla de Kerna, corría del noroeste al sudeste y se cruzaba con la falla de Delfos precisamente en el emplazamiento del oráculo. Una serie de manantiales, que trazaban una línea que recorría el santuario y pasaba por el templo, marcaba la localización de la falla de Kerna bajo antiguas terrazas y detritos acumulados por los desprendimientos de rocas.

Aquel mismo año, el equipo formado por Michael D. Higgins y Reynold Higgins publicaba un libro que nos hizo pensar que íbamos por el buen camino. En su *Guía geológica de Grecia y el Egeo* comentaban que la serie de manantiales indicaba la presencia de una “abrupta falla” que, corriendo en dirección noroeste-sudeste, atravesase el santuario. Apuntaban también que no había ninguna razón geológica que obligara a rechazar la tradición antigua.

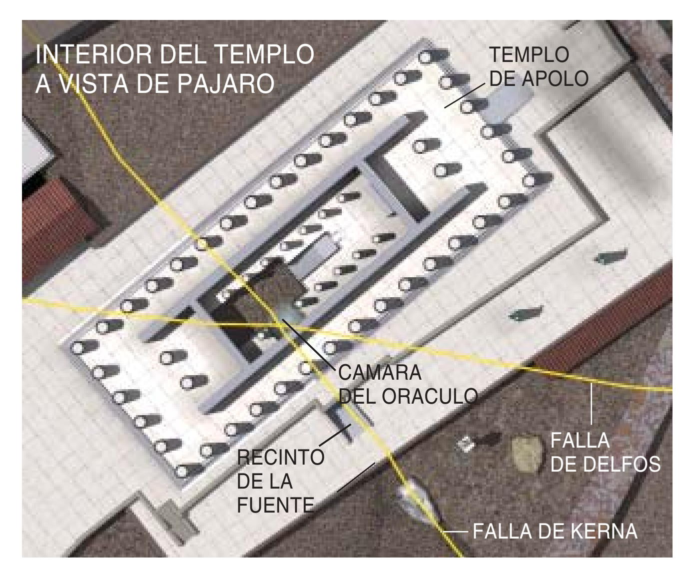
Conjeturaban los Higgins que el gas emitido podría haber sido dióxido de carbono. Una década antes, se había detectado ya una emisión de ese gas en otro templo de Apolo, el de Hierápolis (hoy Pamukkale), en Asia Menor. Siguiendo las indicaciones de Estrabón, los investigadores han descubierto que el templo de Apolo en Hierápolis fue deliberadamente erigido sobre una fumarola que emitía gases tóxicos; emergían de una gruta situada entre los cimientos del edificio.

El templo de Hierápolis no fue un lugar de profecía, pues el dióxido de carbono era, más que embriagador,



4. DOS FALLAS GEOLOGICAS se cortan en Delfos por debajo del templo de Apolo (*arriba, y detalles en la página contigua*). La intersección de las fallas hizo más permeable la piedra y abrió unas grietas (*véaselas en la sección transversal*) por las que pudieron ascender las aguas subterráneas y los gases. La actividad tectónica iría calentando la piedra caliza adyacente a las fallas hasta temperaturas suficientes para vaporizar algunos de sus elementos petroquímicos. Estos vapores gaseosos ascenderían por las grietas abiertas por las fallas y entrarían en la pequeña y cerrada cámara situada a nivel más bajo que el suelo del templo; allí se sentaba la pitonisa para profetizar.

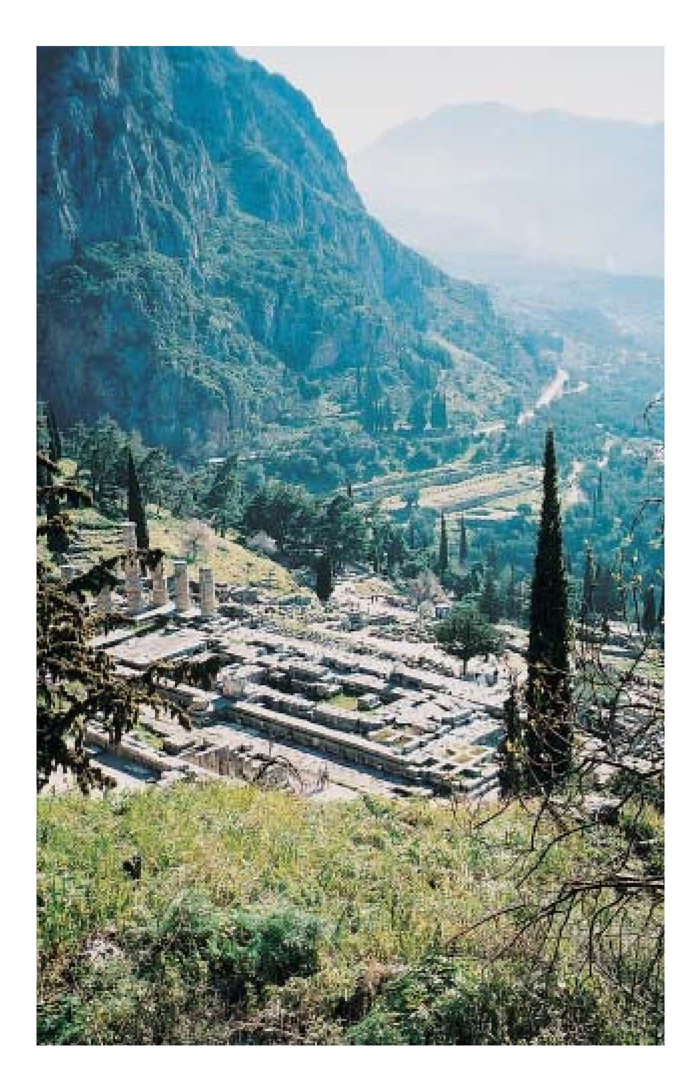
DAVID FIERSTEIN





SECCION TRANSVERSAL
SEGUN EL PLANO
DE LA FALLA DE KERNA
TEMPLO DE APOLO
DEPOSITOS
DE TRAVERTINO
CAMARA DEL ORACULO
RECINTO
DE LA
FUENTE
MANANTIAL
PIEDRA CALIZA
BITUMINOSA
FALLA
DE DELFOS
CARA DE LA FALLA
DE KERNA

DAVID FIERSTEIN; NINA FINKEL (*mapa*)





5. EL TEMPLO DE APOLO (*a la izquierda*), según se ve cuando se mira a lo largo de la falla de Delfos. La cámara oracular se hallaba cerca del bloque que hay en el espacio hundido entre los cimientos del templo. En la fotografía superior, uno de los autores (De Boer) sentado en ese bloque, casi en el mismo sitio en que siglos atrás se sentaría sobre su trípode la pitonisa.

SHANE SOLOW (*izquierda*); JOHN HALE (*derecha*)

venenoso, y quitaba la vida a las víctimas rituales, ya fuesen gorriones o toros. En cambio, otros templos de Apolo que se encuentran en Turquía sí *fueron* oraculares y estuvieron edificados sobre fuentes activas, como el de Dídima y el de Claros. Iba tomando consistencia la idea de la existencia de un vínculo entre los templos de Apolo y los puntos de actividad geológica.

El gas perfecto

Las grietas abiertas bajo el templo por las fallas recién descubiertas en Delfos indicaban por dónde podrían haber subido a la superficie tanto gases como agua de manantial, pero no explicaban el por qué de la generación de los gases. De Boer había observado depósitos de travertino, precipitaciones de calcita producidas por agua de manantial, que revestían las pendientes más arriba del templo e incluso un antiguo muro de contención. Dedujo que el agua fue atravesando gruesos estratos de piedra caliza mientras subía hasta la superficie, donde había depositado las mineralizaciones calcáreas (fenómeno visto también en Hierápolis, Turquía). Varios estudios griegos de la geología del monte Parnaso revelaban que entre las formaciones cretácicas de roca vecinas al templo había capas de piedra caliza bituminosa que contenían hasta un 20 % de compuestos petroquímicos.

De Boer empezaba ya a ver una ilación en el conjunto de datos. Las fallas, bien visibles en las abruptas pendientes del monte Parnaso, habían atravesado la piedra caliza bituminosa. El movimiento de las fallas produjo, con su avance, un rozamiento que fue calentando la caliza hasta un punto en que sus elementos petroquímicos se vaporizaron. Estos salieron entonces a lo largo de las fallas junto con el agua de manantial, sobre todo en los sitios en que el entrecruzarse de las fallas hacía más permeable la roca. Con el tiempo, al ir los trozos de costra calcítica rellenando los huecos del interior de la falla, disminuirían las emisiones de gas, a las que sólo el siguiente movimiento tectónico daría nuevas salidas.

El razonamiento de De Boer parecía concordar con las investigaciones, realizadas a comienzos del siglo XX, por los arqueólogos franceses, quienes, pocos años después de publicarse el artículo de Oppé, habían llegado hasta la roca madre subyacente bajo el *adyton*. Bajo un estrato de arcilla encontraron una roca "agrietada por la acción de las aguas". Nosotros creemos que esas grietas podrían haberse causado por fracturación y creación de fallas, más bien que por la acción del agua, aunque quizá las aguas subterráneas agrandasen con el tiempo las grietas. Creemos también que la hendidura que podía verse en el *adyton* debió de ser una grieta que fue abriéndose hasta alcanzar la capa de arcilla situada sobre el lecho de roca donde se abre la falla.

Los autores

JOHN R. HALE, JELLE ZEILINGA DE BOER, JEFFREY P. CHANTON y HENRY A. SPILLER han formado un equipo interdisciplinar para investigar el oráculo de Delfos. Hale, arqueólogo de la Universidad de Louisville, ha escrito ya dos artículos para *Investigación y Ciencia*. De Boer enseña geología en la Universidad Wesleyana. Chanton, químico, es profesor del departamento de oceanografía de la Universidad estatal de Florida. Spiller, toxicólogo, dirige en Kentucky el Centro Toxicológico Regional.

Nos quedaba por dilucidar qué gases podrían haberse emitido. De Boer supo que unos geólogos que trabajaban en el Golfo de México habían analizado los gases que burbujeaban a lo largo de unas fallas sumergidas; los que emanaban de las fallas activas en esa zona de piedra caliza bituminosa eran metano y etano, hidrocarburos ligeros. ¿Habría ocurrido lo mismo en Delfos?

Para averiguarlo, pedimos que se nos permitiera tomar muestras del agua del manantial de Delfos y del travertino depositado por las antiguas fuentes. Esperábamos descubrir en esta piedra porosa vestigios de los gases que hubieran surgido en la época clásica. En ese momento, Chanton, químico, se unió a nuestro equipo. En las muestras de travertino recogidas por De Boer y Hale encontró metano y etano, este último producto de la descomposición del etileno. Viajó a Grecia para recoger muestras del agua de las fuentes que hay en el templo y en sus alrededores. El análisis reveló la presencia de metano, etano y etileno en el agua de la fuente de Kerna, que mana en el interior del propio santuario. Como el etileno desprende un olor dulzón, la presencia de este gas parecía concordar con lo que decía Plutarco de un gas que olía como un perfume caro.

Para ayudarnos a entender los efectos que tales gases pudieran producir en individuos encerrados en espacios reducidos como el *adyton*, se unió a nuestro proyecto el toxicólogo Spiller. Su trabajo con "*huffers*" —adolescentes que se drogan inhalando los efluvios de sustancias, como las colas o los disolventes, que, por lo general, contienen hidrocarburos ligeros— le había hecho ver algunos paralelismos con el comportamiento de la pitonisa durante el trance.

Spiller halló aún más paralelismos en los experimentos que sobre las propiedades anestésicas del etileno había realizado, tiempo atrás, Isabella Herb. Descubrió que, si bien la mezcla de un 20 % de etileno producía inconsciencia, concentraciones más leves inducían un estado de trance. En la mayoría de los casos el trance era benigno: el paciente permanecía consciente, podía prestar atención y responder a las preguntas, experimentaba sensaciones extracorpóreas y euforia, y sufría amnesia cuando se dejaba de administrarle el gas. Pero en ocasiones vio Herb reacciones violentas: el sujeto lanzaba furiosos alaridos y gritos incoherentes y se agitaba convulso. Si durante tal frenesí algún paciente hubiese vomitado y parte del vómito le hubiese invadido los pulmones, le habrían sobrevenido inevitablemente la neumonía y la muerte. Así pues, de acuerdo con los análisis de Spiller, la inhalación de etileno podría explicar cada una de las diversas características atribuidas al *pneuma* délfico: su dulce aroma, sus variables efectos en sujetos humanos e incluso su poder en ocasiones mortífero.

Una inspiración inesperada

Hace dos mil años se interesó Plutarco por reconciliar la religión con la ciencia. Como sacerdote de Apolo, tenía que responder a lo que algunos conservadores religiosos objetaban, que era inconcebible que un dios se valiese del efluvio de un gas natural para hacer un milagro: ¿por qué no se introducía directamente en el cuerpo de la mujer? Plutarco creía que los dioses, para llevar a cabo sus obras, tenían que contar con los materiales de este mundo corruptible y transitorio. Aunque era un dios, Apolo tenía que anunciar sus profecías mediante las voces de unas mortales a las que había que inspirar con estímulos que formaban parte del mundo natural. Sus cuidadosas observaciones y la meticulosidad con que describieron los gases emitidos en Delfos evidencian que los antiguos nunca trataron de excluir de la comprensión religiosa la investigación científica.

La principal lección que sacamos de nuestro trabajo sobre el oráculo délfico no es el tan socorrido mensaje de que la ciencia moderna puede dilucidar antiguas curiosidades. Quizá tenga más importancia haber advertido lo mucho que ganaremos si abordamos los problemas con la misma amplitud de miras y el mismo talante interdisciplinar de que dieron muestra los griegos.

Bibliografía complementaria

THE DELPHIC ORACLE. H. W. Parke y D. E. W. Wormell. Basil Blackwell, 1956.

PLUTARCH'S MORALIA. Vol. 5. Loeb Classical Library. Sexta impresión. Harvard University Press, 1992.

A GEOLOGICAL COMPANION TO GREECE AND THE AEGEAN. Michael Denis Higgins y Reynold Higgins. Cornell University Press, 1996.

NEW EVIDENCE FOR THE GEOLOGICAL ORIGINS OF THE ANCIENT DELPHIC ORACLE (GREECE). J. Z. De Boer, J. R. Hale y J. Chanton, en *Geology*, vol. 29, n.º 8, págs. 707-711; 2001.

THE DELPHIC ORACLE: A MULTIDISCIPLINARY DEFENSE OF THE GASEOUS VENT THEORY. Henry A. Spiller, John R. Hale, Jelle Z. De Boer, en *Journal of Clinical Toxicology*, vol. 40, n.º 2, págs. 189-196; 2002.

Dinámica de agregados atómicos

El estudio de la química de los agregados en la escala de los femtosegundos proporciona abundante información sobre el instante efímero en que los reactivos se transforman en los productos de la reacción

Lutz Poth, Eric S. Wisniewski y A. Welford Castleman, Jr.

La noción de agregado o cúmulo evoca imágenes que difieren bastante según la disciplina científica de cada uno. El químico estructural puede imaginar las bolas y barras de los modelos moleculares, mientras que a un astrofísico probablemente le vengan a la mente grupos de galaxias espirales. Para los químico-físicos, entre los cuales nos contamos, la palabra "agregado" es quizá más específica: nos habla de una colección de átomos o moléculas débilmente ligados, cuya distribución de tamaños puede modificarse con facilidad por medio de las técnicas con que se forman los propios agregados. Estas pequeñas agrupaciones atómicas tienen propiedades que a veces difieren de manera sustancial de las del material macroscópico; su comportamiento no es el de un cuerpo sólido, un líquido o un gas. A veces se dice que los agregados atómicos constituyen una nueva fase de la materia.

Los agregados son tan pequeños, que una gran parte de sus constituyentes se sitúa en la superficie. Esa pequeñez también impone condiciones a los niveles de energía de los electrones en su interior, de tal forma que a menudo da lugar a efectos cuánticos. La investigación de los agregados apasiona porque aborda nuevos fenómenos fundamentales y tiende un puente entre nuestro conocimiento del estado gaseoso y de la materia condensada gracias a la exploración de estados intermedios. Desde un punto de vista más práctico, podrían servirnos algún día como ladrillos fundamentales para ensamblar nuevos materiales a escala nanométrica.

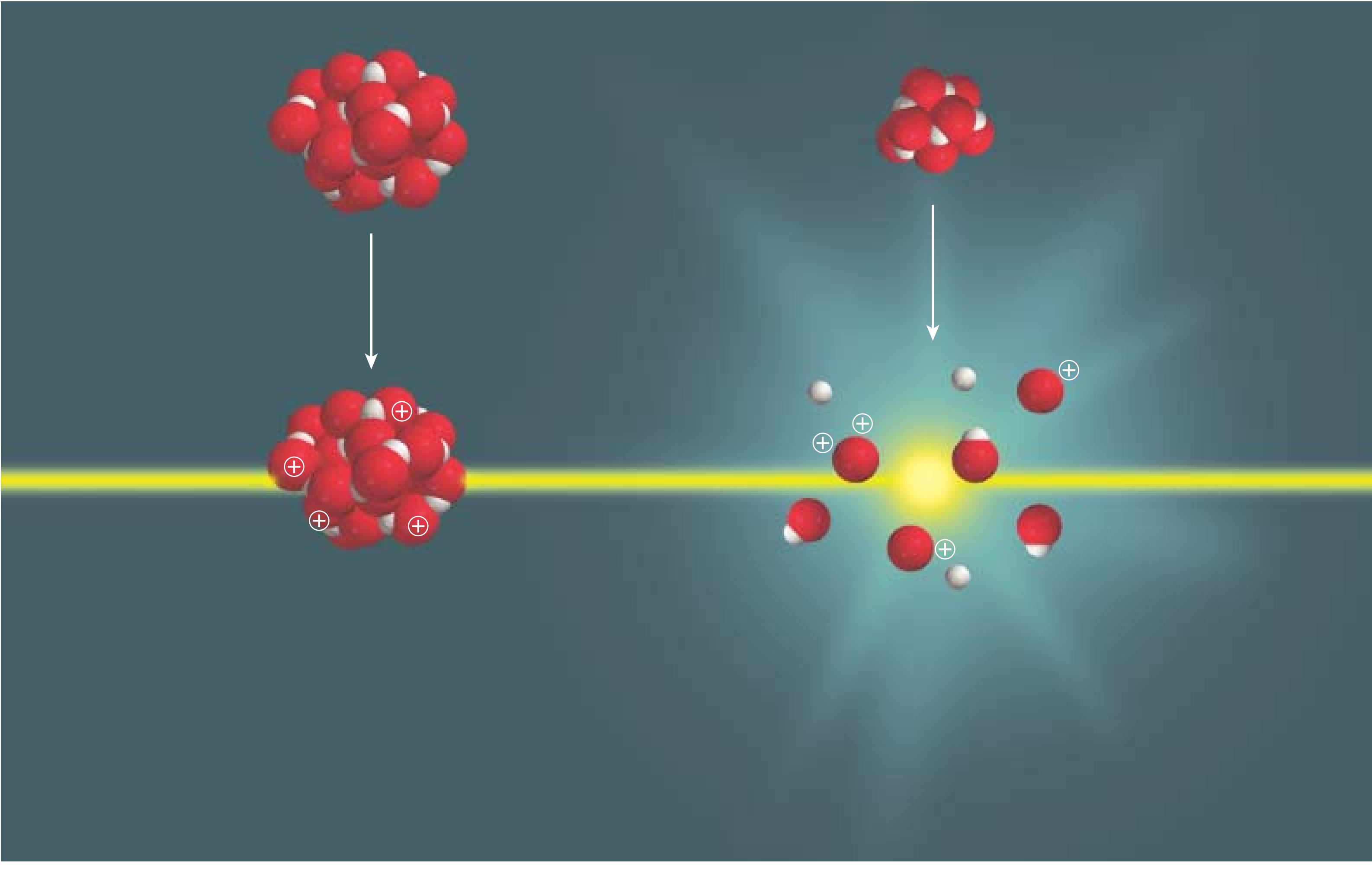
Pero antes de que estas expectativas puedan cumplirse, hay que estudiar más el singular comportamiento físico y químico de los agregados. Y ya que los procesos de interés ocurren en escalas de tiempo cortísimas, se requieren técnicas especiales para una observación directa. En lo que sigue nos centraremos en la dinámica de las reacciones químicas en los agregados, y le dedicaremos particular atención a un fenómeno extraordinario: la explosión coulombiana.

Cargas del mismo signo se repelen

Pocos procesos de interés para la química se efectúan con mayor rapidez que los que involucran especies dotadas de carga. En general, los procesos con transferencia de electrones tienen lugar en las escalas de tiempo más cortas; ahora bien, algunos procesos que discurren entre constituyentes iónicos sólo tardan un poco más en completarse. Un ejemplo de particular importancia para nuestro trabajo es una interacción en la que intervienen átomos y moléculas ionizados. Debido a su carga eléctrica, las partículas ionizadas se influyen a través de distancias comparativamente grandes. Sus interacciones, intensas, a menudo ocurren muy rápido. De la misma manera que se repelen los

El autor

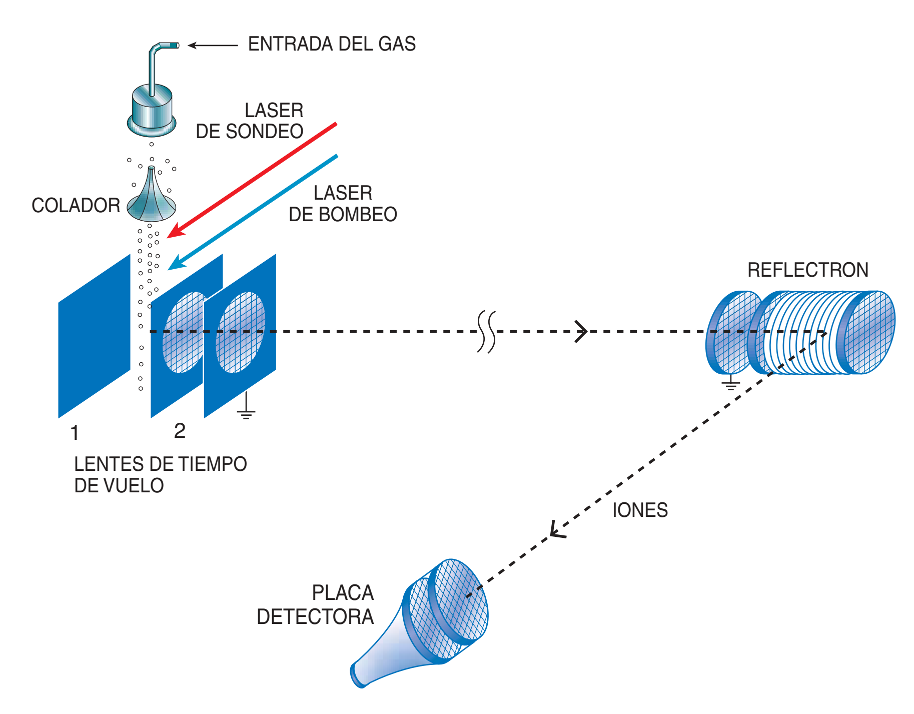
LUTZ POTH trabaja como químico en la División PPC de Degussa Metals Catalysts Cerdec AG, en Frankfurt. Se doctoró en la Escuela Técnica Superior de Darmstadt. Fue después investigador posdoctoral en el Laboratorio de A. W. Castleman. ERIC WISNIEWSKI recibió su doctorado en la Universidad estatal de Pennsylvania, bajo la dirección de Castleman, y es ahora investigador posdoctoral en ese laboratorio. A. WELFORD CASTLEMAN Jr. ocupa la cátedra Eberly de ciencia y la cátedra Evan Pugh de química y física en la Universidad estatal de Pennsylvania. Ha publicado más de 500 artículos.



1. LOS AGREGADOS ATOMICOS son grupos de átomos o moléculas débilmente ligados a los que no puede caracterizarse como sólidos, líquidos o gaseosos. Merced a la facilidad de controlar su composición y tamaño, resultan apropiados para el estudio de los efectos de la solvatación en la dinámica de las reacciones. Tras una rápida ionización producida con una luz láser intensa, los agregados experimentan de inmediato una "explosión coulombiana", una rápida y enérgica repulsión causada por la cercanía entre cargas eléctricas del mismo signo. Los autores, que se valen de la espectrometría de masas por tiempo de vuelo combinada con pulsos láser con una duración de femtosegundos, provocan la explosión coulombiana para detener la evolución del estado de transición durante las reacciones químicas. Aquí se comparan un agregado grande y otro pequeño, ambos sometidos a la ionización por un láser que los despoja de cuatro electrones y les imparte así una carga +4. El agregado más pequeño es incapaz de resistir la repulsión generada por esa ionización y sufre la explosión coulombiana, mientras que el mayor conserva la suficiente fuerza cohesiva y permanece intacto.

polos de los imanes con que juegan los niños, también se repelen entre sí las partículas con igual signo de la carga eléctrica; es la *repulsión de Coulomb*, así denominada en honor a Charles Augustin de Coulomb, el ingeniero y científico francés que propuso las leyes de la atracción y la repulsión eléctricas. Si de repente se genera un conjunto de partículas con el mismo signo de la carga y muy próximas entre sí, la repulsión entre ellas puede resultar tan grande, que haga explotar el sistema: la *explosión coulombiana*. Los agregados atómicos son especialmente propensos a experimentar esos procesos cuando se los somete a los campos luminosos de alta intensidad que proporcionan los láseres de pulsos de femtosegundos. El comportamiento de los agregados durante la explosión coulombiana informa de la evolución en el tiempo de los estados de transición, esas especies de vida efímera que existen transitoriamente durante el desarrollo de una reacción química.

Las consecuencias de la repulsión entre cargas del mismo signo en los agregados se conocen desde hace años, pero las huellas que dejaba el proceso no resultaban muy detalladas. Hasta hace poco, la mayoría de las observaciones de fenómenos coulombianos en agregados atómicos se limitaban a agregados múltiplemente cargados pero intactos, lo bastante grandes además para contener más de un centro de carga; gracias a ello se podía detectarlos con un espectrómetro de masas. En esos casos había suficiente separación entre las cargas como para que no se pudiese vencer la energía de ligadura del sistema; se evitaba de esta forma que el agregado se desmembrase. Pero, por desgracia, las especies que experimentaban la separación explosiva tenían sólo una existencia fugaz y no se podían detectar por medio del espectrómetro. El uso ingenioso de la ionización permitió a veces la observación de un cambio en la estabilidad —debido al calentamiento de los agre-

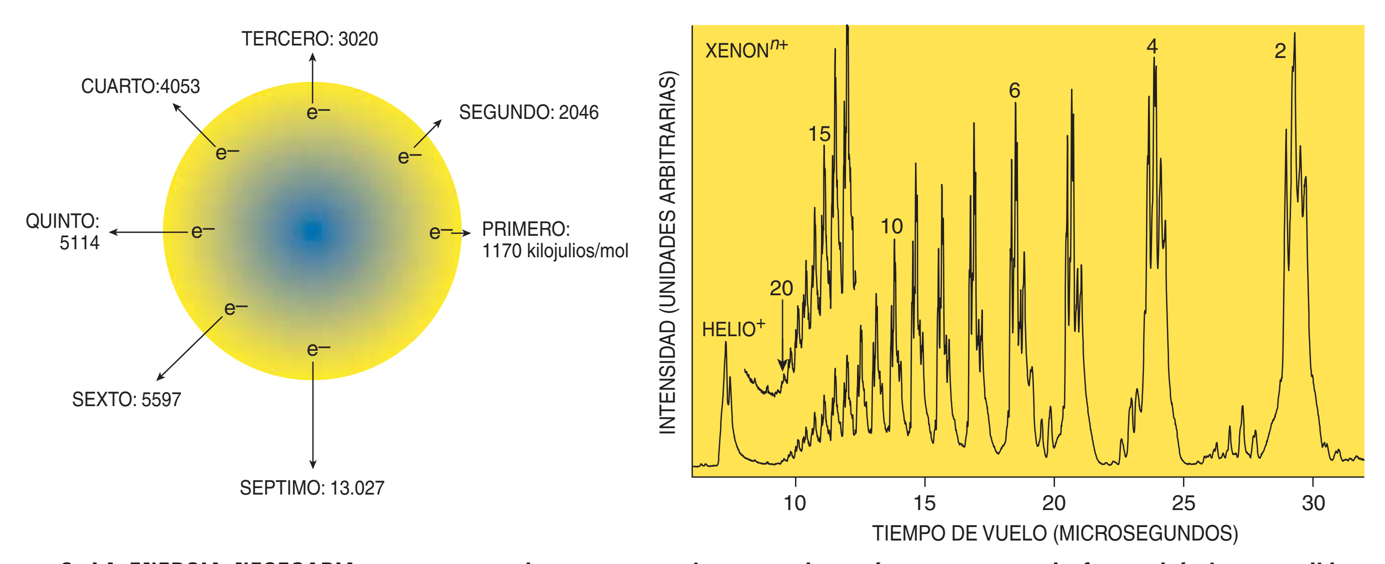


2. LA ESPECTROMETRIA DE MASAS por tiempo de vuelo (TOFMS) mide el tiempo que necesita una partícula cargada para alcanzar un detector. Los agregados que sufren la explosión coulombiana emiten de manera uniforme en todas las direcciones átomos muy cargados. Diseñando su TOFMS de forma que las partículas que se movían en la dirección contraria al detector retrocediesen tras ser reflejadas, los autores midieron el retraso en la llegada de las partículas que recorrían la distancia mayor; es una forma de averiguar la energía cinética. La espectroscopía láser se utiliza también en experimentos de bombeo y sondeo; en ellos, un pulso láser inicia la reacción y un segundo pulso, administrado algunos femtosegundos (o picosegundos) más tarde, "lee" los reactivos y los productos. Ya que el láser de sondeo ioniza los agregados y permite la aceleración de partículas cargadas en el TOFMS, las reacciones terminan en una escala temporal de femtosegundos o picosegundos entre las lentes 1 y 2, aunque la detección ocurre microsegundos más tarde.

gados estables inducido por el láser— que llevaba a su evaporación; la consiguiente reducción del tamaño podía llegar hasta el punto de que fueran tan pequeños como para que la proximidad de las cargas presentes fuera tal, que el sistema no pudiera permanecer estable y se desmembrase. Era posible observar algunos de los efectos coulombianos, pero no la explosión en sí misma.

Observación del interior con un láser

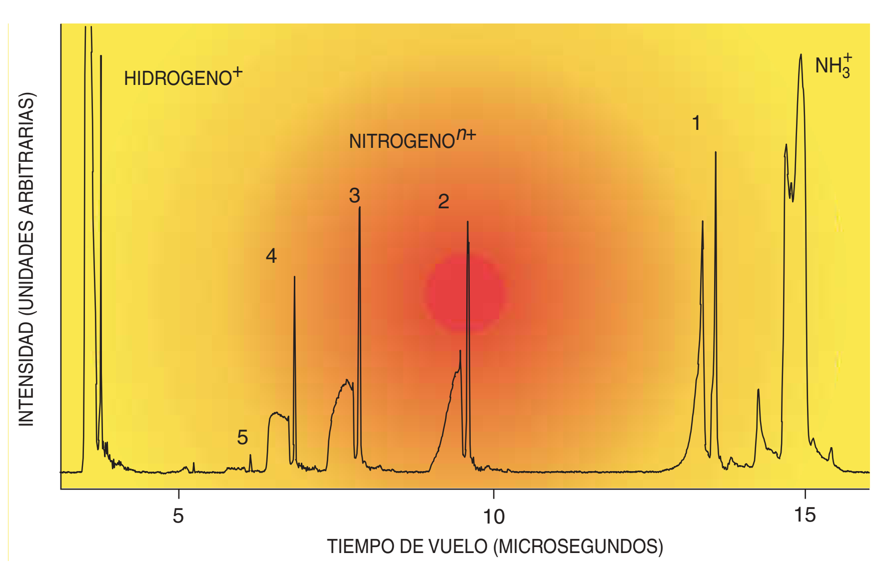
El estudio de las interacciones de átomos, moléculas y materiales sólidos con potentes láseres se ha convertido en una activa línea de investigación desde mediados de la década de 1970. Diversas investigaciones han puesto de manifiesto que las interacciones suficientemente intensas pueden producir fenómenos de ionización múltiple, causados por la expulsión de un gran número de electrones por átomos individuales, aunque el mecanismo preciso no se conoce bien y es un asunto polémico. Hacia la mitad de la década de 1980 se había conseguido producir átomos con carga múltiple, como el helio^{2+} (He^{2+}, un átomo de helio al que le faltan sus dos elec-



3. LA ENERGIA NECESARIA para extraer electrones sucesivamente de un átomo aumenta de forma drástica a medida que se extraen más y más electrones. En el caso del xenón, la extracción del séptimo electrón requiere una energía que es más de un orden de magnitud mayor que la necesaria para quitarle el primero. Los autores han extraído hasta 20 electrones del xenón aplicando pulsos láser intensos y muy cortos. El gas transportador es helio^{+} y los números en la gráfica indican cuántos electrones se extraen.

trones), neón^{2+} y xenón^{5+}, ionizando los correspondientes átomos neutros. Experimentos análogos dieron lugar a argón^{8+}, criptón^{8+}, europio^{6+}, terbio^{5+} e incluso uranio^{10+}. También se crearon átomos multicargados, entre ellos los nitrógenos N^{+}, N^{2+} y N^{4+}, irradiando la molécula N_2 con un láser intenso. La explosión coulombiana de moléculas de yoduro de hidrógeno (IH) ha dado lugar a iones yodo con cargas de hasta +5. Se han estudiado otras moléculas diatómicas, como el monóxido de carbono, oxígeno e hidrógeno. Recientemente la irradiación de la molécula de hexafluoruro de uranio (UF_6) dio lugar a U^{+}, U^{2+}, y U^{3+}.

Hace unos años observamos un fenómeno inesperado: pulsos láser de alta intensidad que creaban al instante, por medio de flujos de luz, sorprendentemente bajos, agregados en un estado muy cargado. La clave para obtener este resultado fue el uso de pulsos láser ultracortos, con una duración del orden de los femtosegundos (un femtosegundo es igual a 10^{-15} segundos). Para hacernos una idea de esta escala de tiempo, imaginemos 25 campos de fútbol, cada uno cubierto, en toda su extensión, con arena de playa hasta una altura de un metro. Si el total de los granos de arena hace las veces de un segundo, un femtosegundo será el equivalente a un grano. Al lanzar un intenso pulso que sólo dure unos femtosegundos contra un agregado se producirá casi instantáneamente la pérdida de muchos electrones; por tanto, se creará un agregado muy cargado, que explotará enseguida. Los mecanismos detallados que llevan a este extraordinario fenómeno son en la actualidad objeto de enorme interés tanto por parte de los científicos teóricos como de los experimentales.



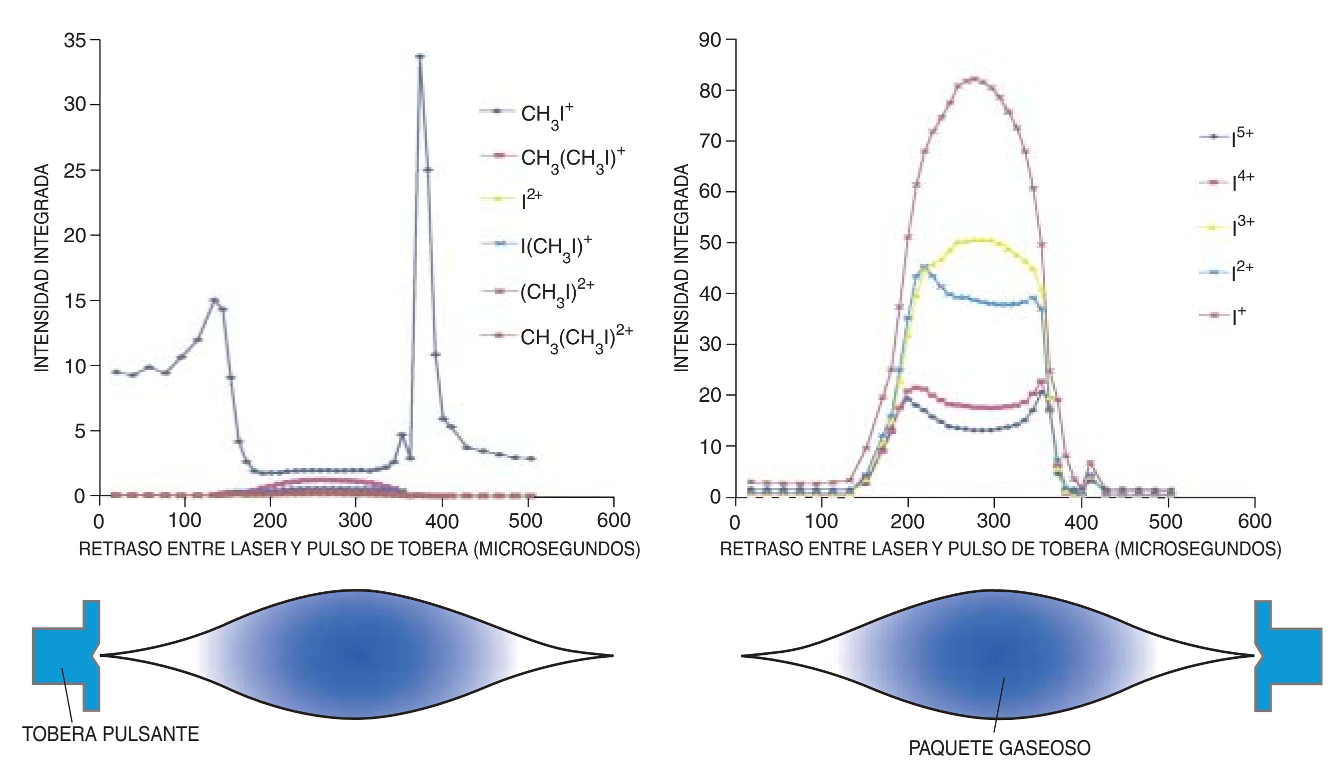
4. LOS MONOMEROS y agregados de amoníaco sufren la explosión coulombiana al ser sometidos a una intensa luz láser, pero sólo los agregados dan lugar a átomos de nitrógeno múltiplemente cargados (con tiempos de llegada de 5 a 10 microsegundos). El monómero (con tiempo de llegada de unos 15 microsegundos) resulta ionizado, pero no se fragmenta en nitrógeno e hidrógeno. Los números indican el grado de ionización del nitrógeno.

Explosión coulombiana en agregados

A medida que ha crecido el interés en el estudio de los agregados atómicos mediante el uso de pulsos de femtosegundos, los nuevos experimentos han ido mostrando que el comportamiento temporal de las moléculas cuando forman parte de un agregado difiere bastante de cuando se hallan libres. Las investigaciones se centran en la irradiación con luz láser intensa de haces de agregados atómicos y moleculares. Cuando la explosión coulombiana hace añicos estos agregados, los estados de carga de los fragmentos atómicos y moleculares pueden variar; su energía cinética llega incluso a ser de varios miles de electronvolt (eV); en algunos casos de sistemas muy cargados, incluso de millones de electronvolt. Con los elevados flujos láser que se aplican a menudo, la explosión coulombiana va a veces acompañada de la emisión de rayos X.

La investigación de las raíces de la explosión coulombiana en agregados recurre en ocasiones al espectrómetro de masas por tiempo de vuelo (TOFMS), que mide el tiempo que tarda una partícula ionizada en llegar a un detector situado a una distancia fija del lugar en el que se crea el ion (*véase la figura 2*). Los agregados que experimentan la explosión coulombiana emiten uniformemente en todas las direcciones átomos muy cargados, aunque sólo se detectan los despedidos en la dirección que corre a lo largo del eje del espectrómetro de masas. Situábamos el TOFMS de tal modo que las partículas que saliesen en dirección contraria a él se reflejasen eléctricamente y volvieran a él. Ya que estas partículas recorrerán un camino mayor que las emitidas por el lado del agregado que mira hacia el detector, llegarán más tarde a éste. Esa diferencia en el tiempo de llegada es una medida de la fuerza repulsiva en el momento de la explosión y vale para calcular la energía cinética asociada con el suceso. Además, el espectro de masas presenta las explosiones coulombianas como picos diferenciados y ligados a las especies que aparecen en los fragmentos. Los iones emitidos hacia adelante por la explosión engendran un pico ancho con un tiempo de llegada corto, mientras que los iones emitidos en la dirección contraria al detector llegan cierto tiempo después y dan lugar a estrechos picos individuales (*véase la figura 4*).

Los espectros de masas de las moléculas de amoníaco ionizadas con láseres de distinta potencia muestran sólo moléculas ionizadas (NH_3^{+}) cuando la potencia del láser se mantiene por debajo del umbral a partir del cual se emite un gran número de electrones. Al ir aumentando la potencia, van apareciendo algunos átomos de nitrógeno múltiple-



5. CUANDO SE INYECTAN en el espectrómetro, pulso a pulso, paquetes de yoduro de metilo espaciados regularmente y sometidos a una explosión coulombiana, sólo los agregados producen especies muy ionizadas. Ya que los agregados se forman sobre todo en el centro del paquete, deberíamos esperar la presencia de los fragmentos ionizados típicos de la explosión coulombiana sólo en la zona media de cada pulso; eso es justo lo que aparece en la gráfica de la derecha. Los espectros de los paquetes moleculares de yoduro de metilo indican la presencia de especies moleculares ionizadas (*izquierda*) al principio y al final del pulso.

mente cargados, así como iones H^+. Por el contrario, si en vez de monómeros se ionizan agregados de amoníaco mediante intensos haces de láser, la producción de átomos de nitrógeno multicargados aumenta considerablemente y se libera una gran cantidad de energía cinética. Además, la explosión coulombiana produce agregados ionizados intactos; los experimentos que emplean la organización covariante (una técnica estadística que pone en concordancia fenómenos físicos relativos a los productos para identificar un origen común) correlacionan esos agregados con un progenitor específico.

Lo mismo que en el caso de los agregados de amoníaco, los agregados de IH, así como los heteroagregados de IH y argón, irradiados con pulsos láser de femtosegundos, producen átomos de yodo en estados de carga nada menos que de I^{17+} y átomos de argón en estados de carga hasta de Ar^{8+}: sorprendentemente, estos átomos pierden todos sus electrones de valencia. Mediante la organización covariante se llega a la conclusión de que los I^{n+} tienen su origen en la explosión coulombiana de agregados de IH, mientras que los Ar^{n+} surgen en la explosión de los agregados $IHAr_m$. La liberación de una energía cinética de varios cientos de eV acompaña a estas reacciones.

Otro método directo para investigar el papel de los agregados consiste en variar de forma selectiva el momento en que el rayo láser se cruza con el haz molecular de agregados. Ya que sabemos que los componentes, tanto al principio como al final de la expansión, son en su mayor parte especies no agregadas, podemos buscar una correlación entre las partículas dotadas de carga y la presencia de agregados durante los estadios intermedios de la expansión. Para flujos de láser moderados, la mayoría de las especies cargadas están presentes cuando, y sólo cuando, los agregados son la especie dominante en el haz molecular (*véase la figura 5*). Todos los indicios apuntan a una conexión directa entre agregados y fragmentos muy cargados y con alta energía cinética.

¿Por qué se dan cargas tan grandes?

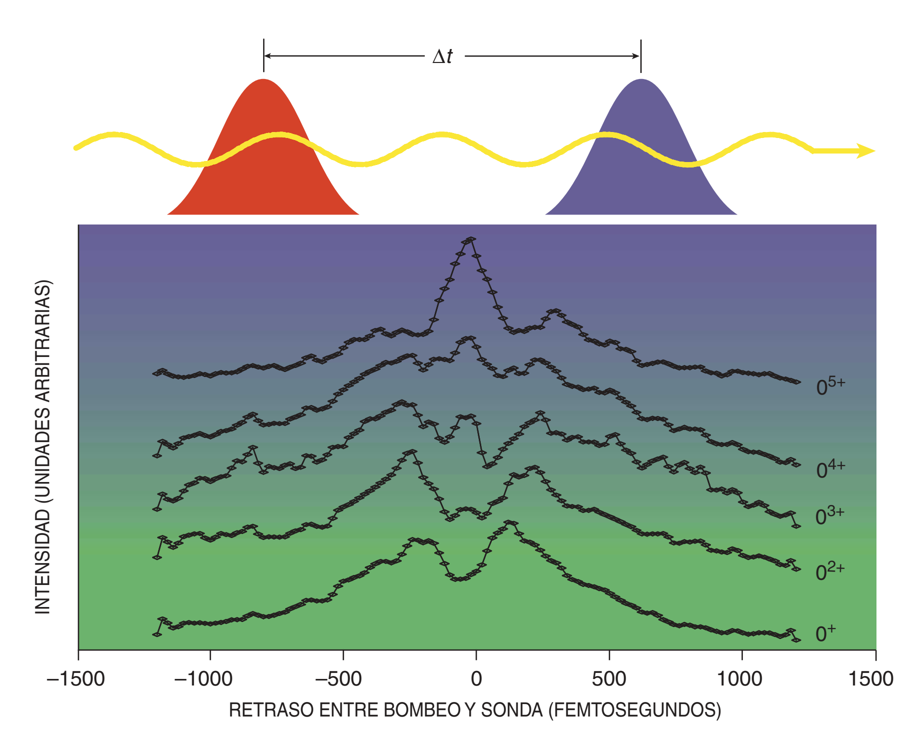
Se han propuesto dos mecanismos de ionización para explicar la formación de fragmentos multicargados cuando se irradian las moléculas o los agregados con intensos haces de luz. Según el "modelo de movimiento electrónico coherente" (MMEC), los electrones ionizados por efecto del campo se mueven de manera coherente; aumenta así el número de ionizaciones a medida que los electrones empiezan a comportarse conjuntamente como cuasipartículas. En términos generales, la formación de iones muy cargados se entiende en este modelo como un caso de ionización por impacto

electrónico. Una explicación alternativa, el "modelo de ionización por ignición" (MII), afirma que los estados altamente ionizados se deben a la disminución del umbral de ionización causada por la presencia cercana de centros iónicos creados durante las ionizaciones iniciales. El modelo MII se basa en consideraciones muy detalladas acerca de la dinámica de los movimientos electrónicos inducidos por ionización por resonancia de la carga (IREC). Conforme al fenómeno de la IREC, se predice que la tasa de ionización depende de una manera no lineal de la distancia internuclear; supone que la ionización produce una pérdida secuencial de electrones que se extiende sobre distancias internucleares cada vez mayores. Una modificación del mecanismo IREC, la IREC dinámica, tiene en cuenta el comportamiento transitorio del pulso láser y su acoplamiento con el movimiento de los núcleos del agregado.

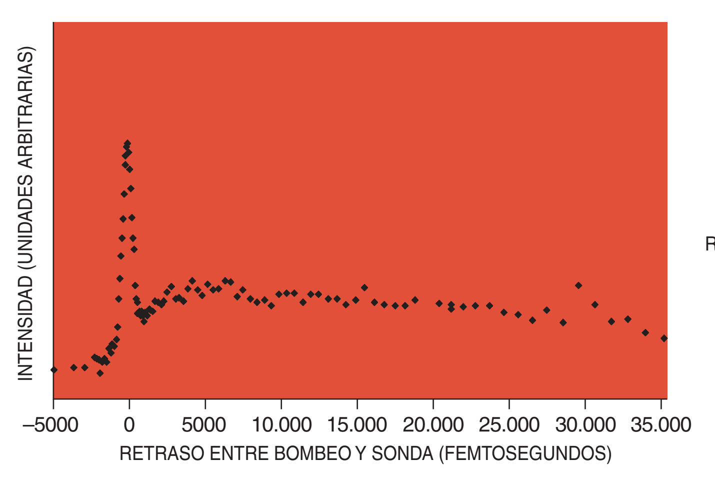
Realizamos en nuestro laboratorio unos experimentos con agregados de acetona. En éstos, un pulso de determinado color del láser de femtosegundos excita el sistema y un pulso de otro color, aplicado un instante después, explora su espectro de masas (a este tipo de experimento se le denomina de "bombeo y sondeo"). Observamos las pronosticadas oscilaciones temporales en la producción de átomos de carbono y oxígeno debidas a la explosión coulombiana. Este hallazgo respalda en general el mecanismo IREC y, por tanto, el MII. Otros elementos observacionales de juicio, sobre todo las energías cinéticas de los fragmentos multicargados resultantes (de decenas o centenas de eV), respaldan, sin embargo, a ambos mecanismos: MII y MMEC. Muchos investigadores piensan que el modelo MII/IREC opera en agregados pequeños y el MMEC es válido para partículas mayores, de tamaño comparable a la longitud de onda de la luz empleada.

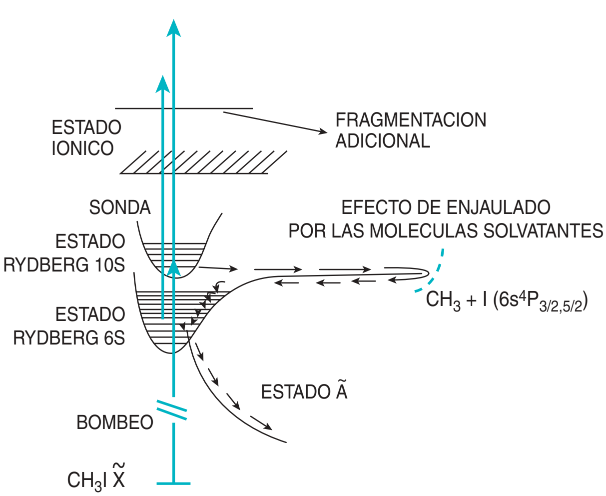
En los experimentos de bombeo y sondeo con agregados de acetona se crean fragmentos de oxígeno transitorios que han perdido de uno a cinco electrones. En la gráfica de intensidades se ve una caída notable en el estado de transición O^+ cuando el retraso es nulo; para especies sucesivamente más cargadas aparece un pico creciente. Cuando la pérdida es de 2 o 3 electrones, la señal iónica cae a un mínimo local, y vuelve a un máximo local para cierto retraso entre bombeo y sondeo. Para retrasos sustanciales, las pérdidas de 2 a 5 electrones presentan un patrón de "latidos"; demuestra que es posible obtener un control cuántico por medio de láseres de alta intensidad. Los máximos de O^+ y O^{5+} se encuentran en la región de tiempos de demora positivos, mientras que los de O^{2+} hasta O^{4+} caen en la de tiempos negativos. Hemos descubierto que los fragmentos formados por carbono con niveles de pérdida de 1 a 4 electrones se comportan de manera análoga. Esas variaciones en la ionización según el retardo entre bombeo y sondeo coinciden con lo esperado conforme al modelo MII/IREC.

Estas oscilaciones de la intensidad merecen mayor comentario. En primer lugar, los máximos y los mínimos no ocurren para unas mismas demoras temporales en las diferentes especies O^{n+}; semejante respues-



6. EN UN EXPERIMENTO DE BOMBEO Y SONDEO, las especies detectadas por el segundo pulso láser (el de bombeo) dependen de la magnitud del retardo —y por tanto del progreso de la reacción— entre los dos pulsos. Aquí se representa la presencia de átomos de oxígeno multicargados obtenidos a partir de agregados de acetona que experimentan la explosión coulombiana (el perfil gaussiano *azul* representa el pulso del láser de bombeo y el *rojo* el pulso sonda). Dos aspectos de los patrones mostrados favorecen el modelo de ionización por ignición (MII), que intenta explicar la extraordinaria presencia de especies muy cargadas después de la explosión coulombiana. En primer lugar, la pronunciada depresión de O^+ para una demora cero y el cambio gradual con el aumento de la ionización hasta llegar a un estrecho pico en O^{5+}, dan a entender que el incremento de las especies muy ionizadas aumenta aún más la ionización, a causa de la disminución del umbral de ionización provocado por los centros iónicos vecinos. En segundo lugar, se percibe un notable patrón en forma de "latidos", en particular para las especies más cargadas. Puede explicarlo la ionización por resonancia de carga, que predice que la tasa de ionización depende de una manera no lineal de las distancias entre los núcleos. Según sea la demora, la distancia entre los núcleos aumentará cierta cantidad, con la consiguiente modificación de la función de onda y su localización; de ahí resultan diferentes tasas de ionización para los distintos estados de carga.





7. MEDIANTE LA VARIACION DE LA DEMORA del pulso sonda es posible elaborar una "imagen" del desarrollo de una reacción. Lo que "se ve" puede ser mucho más complejo de lo que la concepción ordinaria de las reacciones químicas haría esperar. La figura muestra la evolución del yoduro de metilo durante el proceso de bombeo y sondeo. Véanse los efectos transitorios en los cambios de la intensidad con el transcurso del tiempo. Pueden explicarse por medio del diagrama de estados energéticos de la derecha. Tras ser excitado al estado de Rydberg 10s, el yoduro de metilo se disocia, pero como los átomos permanecen muy próximos por un efecto llamado de *enjaulado*, se recombinan en el estado Rydberg 6s. Este producto de la recombinación sufre la desexcitación mediante relajación vibracional interna hasta que alcanza la superficie de cruce con el estado disociativo Ã y se disocia de nuevo.

ta a diferentes retrasos se repite en todas las medidas tomadas por el método de bombeo-sondeo. No son una casualidad. La pauta no puede ser un efecto espurio de la fase de luz láser de bombeo o de sondeo, ya que el período correspondiente a su longitud de onda es de sólo 2 femtosegundos. Además, la potencia de los láseres de bombeo y sondeo basta para provocar la ionización múltiple de los agregados de acetona y, por lo tanto, es improbable que el agregado pueda estabilizar la gran cantidad de energía potencial correspondiente de forma que se cause una disminución de la intensidad. Estadísticamente, tampoco encontramos ninguna periodicidad en los estados de transición. Tengamos en cuenta, sin embargo, la predicción del modelo IREC: en razón de la demora temporal del haz de sondeo, los espaciados interatómicos crecerán hasta cierta distancia como resultado de los movimientos nucleares debidos a la explosión coulombiana. A medida que las distancias interatómicas aumentan, las funciones de onda y, por tanto la localización electrónica, cambian, lo que lleva a tasas de ionización diferentes para los distintos estados de carga. Esto daría lugar a los máximos y mínimos observados en los estados de transición.

Dinámica de reacción en agregados

Todos los libros de texto de química explican de manera similar las reacciones químicas entre los reactivos que engendran los productos de reacción. Se centran en la cinética de la reacción, que se observa siguiendo la evolución temporal de la concentración cambiante de reactivos y productos. Aunque esto proporciona una valiosa información, nos dice muy poco acerca de los mecanismos moleculares de solvatación y disociación, e ignora los importantes estados intermedios. Para entender las reacciones químicas hay que observar cómo se van rompiendo y formando enlaces en las pertinentes escalas de tiempo. Sólo es posible desde hace unos diez años.

La velocidad y la escala de las reacciones químicas convierten ese objetivo en una tarea formidable. Los átomos se mueven y forman enlaces entre sí a velocidades que suelen caer fuera de nuestro alcance: un billón de veces más rápido que el parpadeo de un ojo. Con unas longitudes de enlace del orden del nanómetro (10^{-9} metros) y velocidades nucleares de alrededor de 1000 metros por segundo, la resolución temporal de los instrumentos ha de ser del orden de unos pocos femtosegundos.

Gracias a la introducción de las técnicas ultrarrápidas de láser en la óptica física, se han podido construir herramientas que posibilitan la exploración directa en esas escalas temporales. La femtoquímica, de la que fueron pioneros Ahmed Zewail, premio Nobel de química en 1999, y sus colaboradores, inicia las reacciones químicas por medio de fotones de bombeo con longitudes de onda escogidas para proporcionar la energía específica que se necesite. El transcurso de la reacción se examina entonces por fluorescencia o por ionización; por sonda se utilizan pulsos de láser convenientemente retrasados.

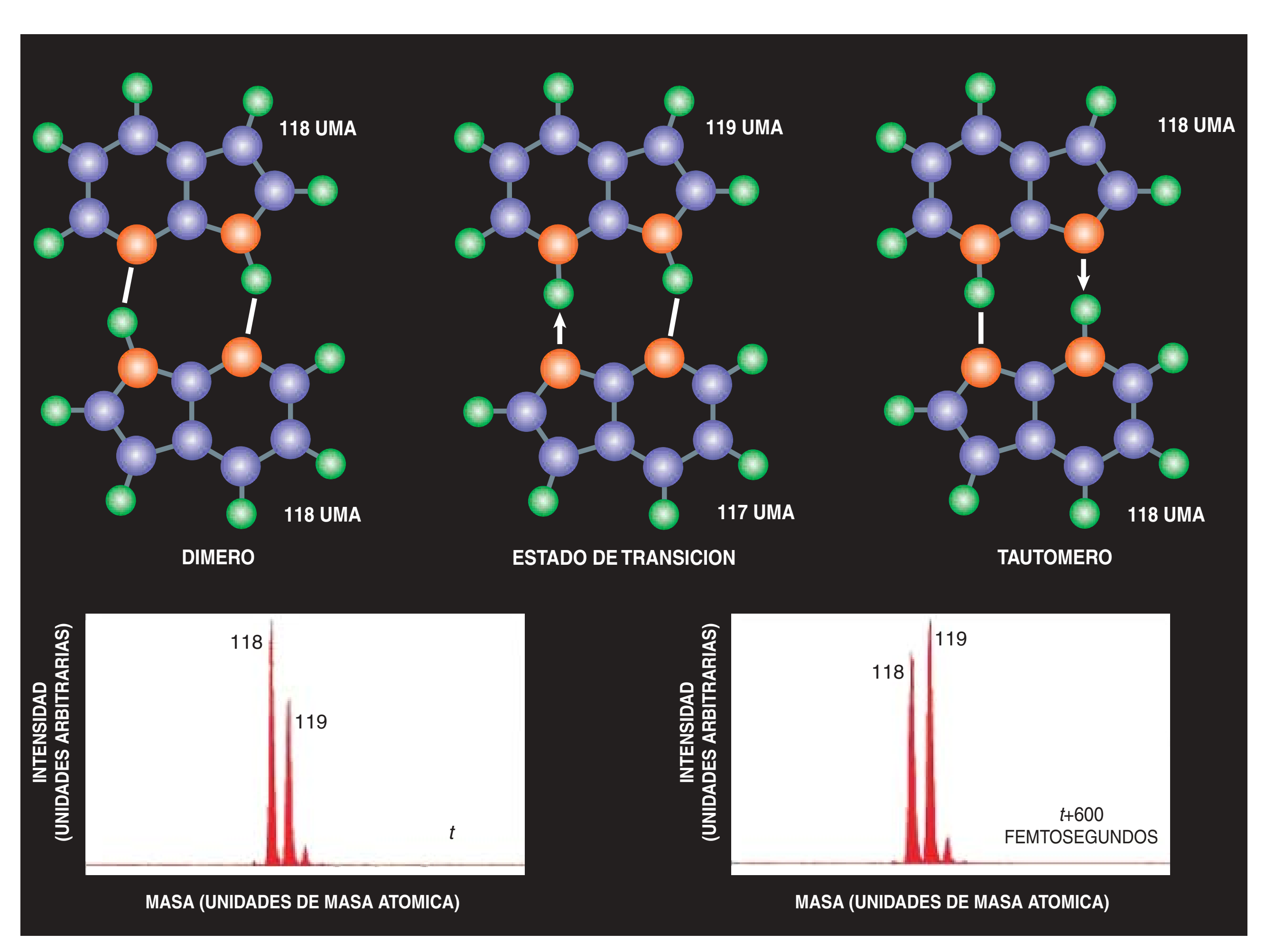
La femtoquímica de agregados presenta un particular interés porque sus estados de transición son a menudo bastante largos y permiten la observación de las consecuencias de la solvatación sobre una transformación reactiva. Tomemos como ejemplo el yoduro de metilo. Comparemos una molécula aislada y un agregado (en que dos moléculas se unen y crean un dímero), excitados la una y el otro hasta el mismo estado electrónico mediante un fotón de bombeo. Obtendremos que los estados de transición de los

iones de yoduro de metilo producidos a partir del agregado exhiben significativas diferencias con respecto a los generados cuando sólo hay monómeros en el haz molecular (*véase la figura 5*). Al sondear con distintos retardos "veremos" lo que ocurre.

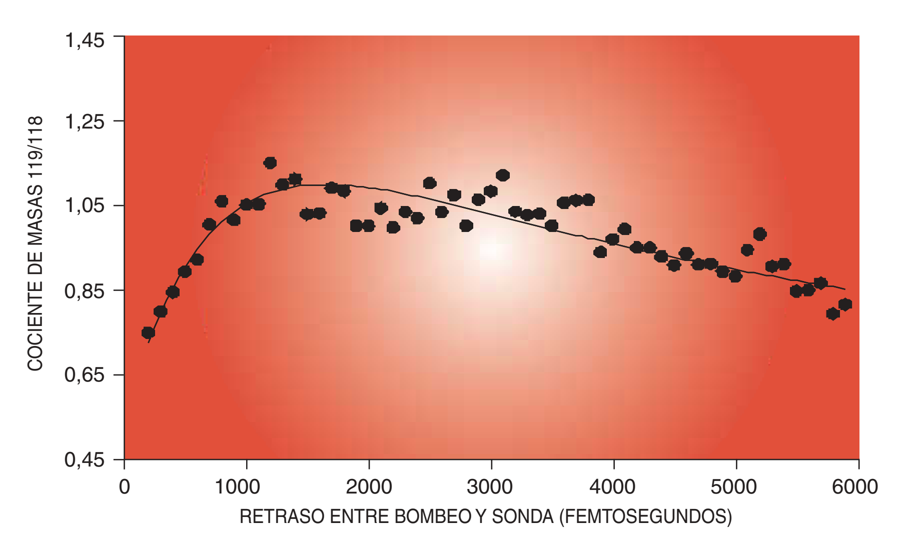
Las observaciones experimentales de ese sondeo (*véase la figura 7*) son compatibles con un mecanismo en el cual se alcanza un estado del yoduro de metilo, el Rydberg 10s, mediante dos fotones de bombeo. El aumento de la vida media del estado Rydberg 10s en el agregado puede explicarse por un cambio de la energía inducido por el agregado y una reducción del acoplamiento de este estado y el estado disociativo (el estado Ã), efecto causado por la moléculas vecinas en el agregado. La señal se recupera pasados 3 picosegundos porque el estado Rydberg 10s acaba por disociarse y se forman átomos de yodo ionizados. Estos se mantienen próximos debido a un efecto de "enjaulado", en el que se intercambia energía entre moléculas cercanas. Los átomos de yodo se recombinan para formar yoduro de metilo en un estado Rydberg más bajo; la ionización desde este estado Rydberg conduce a un aumento en la señal iónica para tiempos de demora mayores de 60 picosegundos. Finalmente, la molécula de yoduro de metilo experimenta relajación vibracional y se disocia de nuevo a través del estado Ã.

Detención de los estados intermedios

Nuestro interés por la explosión coulombiana iba más allá de la explicación de sus mecanismos básicos. Alcanzaba a sus posibles



8. SEGUN UNA HIPOTESIS DE WATSON Y CRICK, una transferencia protónica desfavorable podría hacer que los pares de bases del ADN fueran susceptibles de sufrir mutaciones puntuales causadas por desajustes de los tautómeros. Para que esto ocurra ha de existir, sin embargo, un par de bases modificado, que permita una discordancia en el emparejamiento. Las figuras de bolas y barras muestran cómo podría acontecer una tautomerización *por pasos*. El grupo de Ahmed Zewail fue el primero en obtener pruebas de que la reacción avanza por pasos; como sistema modelo se valieron del 7-azaindol. Los autores han profundizado esa investigación usando la explosión coulombiana para detener la reacción y detectar directamente el estado de transición con 119 unidades de masa atómica (uma) que se crea en la tautomerización del 7-azaindol. Las dos gráficas muestran las proporciones de monómeros con 118 y 119 uma en un instante inicial y 600 femtosegundos después de iniciada la reacción. El pico de 119 uma en la gráfica de la derecha revela que hay un estado de transición y que la reacción procede por pasos.



9. LA EXPLOSION COULOMBIANA del dímero de 7-azaindol revela dos marcas cinéticas características por medio de la proporción entre el monómero de 119 uma y el de 118 uma. La primera, a los 600 femtosegundos, indica la transferencia del primer protón. La segunda, a unos 3000 femtosegundos, señala la transferencia del segundo. El uso de la explosión coulombiana, que separa las especies en una escala temporal de sólo 20 femtosegundos, hace posible la detección de estas especies intermedias.

aplicaciones. En particular, comprendimos que la explosión coulombiana, aplicada selectivamente, podría usarse para estudiar la evolución en tiempo real de los productos intermedios en una reacción rápida. Nos hemos valido de técnicas de bombeo y sondeo; con un pulso sonda de alta energía generamos la explosión coulombiana del estado de transición de la reacción para determinados tiempos de demora, una vez se ha iniciado la reacción. Los estados de transición así detenidos se detectan directamente con un TOFMS; de ese modo se explora y "visualiza" la dinámica de la reacción.

Una aplicación de especial interés se remonta a una hipótesis de Watson y Crick: que una transferencia protónica desfavorable entre pares de bases del ADN pudiera hacerlas susceptibles a mutaciones puntuales causadas por desajustes entre los *tautómeros* (dos o más isómeros —moléculas que sólo difieren en su configuración espacial— que pueden convertirse unos en otros de manera que sus proporciones se mantengan en equilibrio). Puede ocurrir si la transferencia protónica crea pares de bases modificados de manera que aumente la probabilidad de un apareamiento inapropiado. El grupo de Zewail aportó una prueba clara de la existencia de más de un paso en un proceso de transferencia protónica; lo lograron sometiendo al bombeo-sondeo un sistema modelo, el dímero 7-azaindol. Si el proceso consta de más de un paso, se transferirá primero un protón de un monómero del dímero al otro. Se creará así un estado intermedio en el cual el monómero que recibe la transferencia tendrá ligados dos protones más —119 unidades de masa atómica en total— que el otro monómero —que se quedará con 117 unidades—. Crecerá así la posibilidad de errores en el emparejamiento. En el segundo paso de la reacción, un segundo protón se trasladará desde la especie de masa 119 hasta la especie de masa 117 y habrá de nuevo dos componentes de masa 118. Si, por otra parte, el proceso es *concertado*, es decir, si los protones se transfieren a la vez, no se producirá ningún estado de transición en la formación del tautómero.

El estado de transición sólo se puede detectar si se detiene la reacción con una explosión coulombiana, que separará las especies individuales en una escala temporal pequeñísima (unos 20 femtosegundos). Indujimos, pues, explosiones. Observamos que la proporción de fragmentos de masa 119 y masa 118 cambiaba al variar el tiempo de retraso entre bombeo y sondeo. Quedó claro que la tautomerización se realiza en este sistema modelo en más de un paso. Los tiempos característicos medidos para las dos transferencias de protones son de unos 600 femtosegundos y 5 picosegundos (5000 femtosegundos), respectivamente (*véase la figura 9*). Nos atrevemos a anticipar que este método de detección de los estados de transición y de examen de su estructura en distintos instantes se va a convertir en un procedimiento útil para estudiar la dinámica de una amplia familia de reacciones; tendrá sus ventajas con respecto a los métodos tradicionales de espectroscopía ultrarrápida.

Perspectivas de futuro

Recientemente, se han llevado a cabo experimentos en el Laboratorio Nacional de Livermore encaminados a generar la fusión nuclear en agregados de deuterio por medio de la explosión coulombiana. Este método pudiera encontrar uso como fuente de neutrones; los aparatos serían de tamaño reducido. Con la explosión coulombiana se ha desentrañado la estructura completa de moléculas pequeñas. El ion investigado, por ejemplo el CH_4^+, se acelera en una máquina Van der Graaf muy potente hasta una energía cinética de unos 200.000 eV por unidad de masa atómica. Entonces se le hace incidir sobre un blanco, una lámina delgadísima, donde los iones pierden los electrones responsables del enlace molecular en centésimas de femtosegundo. Los productos de la explosión coulombiana resultante atraviesan el blanco y se registran con un detector sensible a la posición y al tiempo. Los datos sirven para deducir la estructura de la molécula inicial.

Es de suponer que los agregados pequeños podrían explorarse de una forma similar; como su propensión a experimentar la explosión coulombiana deberían permitir que las energías requeridas fuesen menores, esas investigaciones estarían al al-

cance de numerosos laboratorios. Con el progresivo desarrollo de los láseres de estado sólido del tipo Ti:zafiro, que caben en una mesa y cuyos pulsos cada vez son más breves y de mayor energía, resulta fácil su aplicación práctica a interesantes e importantes problemas de química, física y biología; así, tanto los expertos en los láseres como los que no lo son pueden entrar en este campo fascinante. La excitación de especies químicas por láser hasta estados electrónicos bien definidos debería enseñarnos mucho sobre la configuración inicial de los núcleos, ya que los procesos de pérdida electrónica tienen lugar en una escala temporal mucho más corta que la de los movimientos de los núcleos. A medida que resulte más fácil sondear la estructura y la dinámica de las reacciones de los agregados, esperemos que esta capacidad se aplique en muchos campos.

Bibliografía complementaria

ARRESTING INTERMEDIATE STATES IN A CHEMICAL REACTION ON A FEMTOSECOND TIME SCALE: PROTON TRANSFER IN MODEL BASE PAIRS. D. E. Folmer, L. Poth, E. S. Wisniewski y A. W. Castleman, Jr. en *Chemical Physics Letters*, vol. 287, pág. 1; 1998.

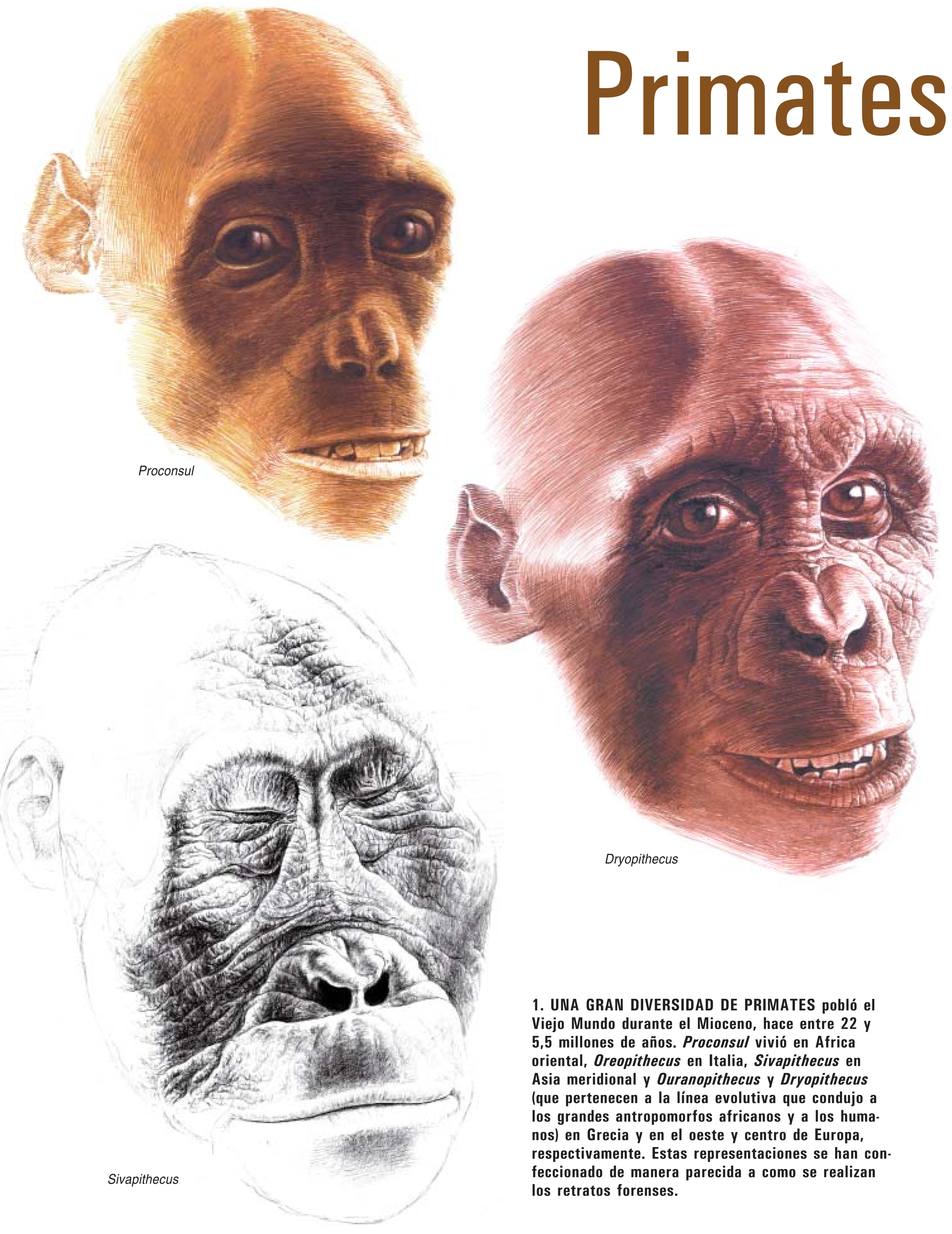
FEMTOSECOND CLUSTER STUDIES OF THE SOLVATED 7-AZAINDOLE EXCITED STATE DOUBLE-PROTON TRANSFER. D. E. Folmer, E. S. Wisniewski, S. M. Hurley y A. W. Castleman, Jr. en *Proceedings of the National Academy of Sciences*, vol. 96, págs. 12.980. 1999.

FEMTOSECOND LASER INTERACTIONS WITH METHYL IODIDE CLUSTERS. J. V. Ford, Q. Zhong, L. Poth, y A. W. Castleman, Jr. en *Journal of Chemical Physics*, vol. 110, págs. 6257-6267; 1999.

FEMTOSECOND LASER INTERACTIONS WITH METHYL IODIDE CLUSTERS. 2: Coulomb Explosion at 397 nm. J. V. Ford, Q. Zhong, L. Poth, y A. W. Castleman, Jr. en *International Journal of Mass Spectrometry and Ion Processes*, vol. 192, págs. 327-345; 1999.

FEMTOCHEMISTRY AND FEMTOBIOLOGY: ULTRAFAST DYNAMICS IN MOLECULAR SCIENCE. Dirigido por A. Douhal y J. Santamaría. World Scientific Publishing Company; Singapur, 2002.

Primates



1. UNA GRAN DIVERSIDAD DE PRIMATES pobló el Viejo Mundo durante el Mioceno, hace entre 22 y 5,5 millones de años. *Proconsul* vivió en Africa oriental, *Oreopithecus* en Italia, *Sivapithecus* en Asia meridional y *Ouranopithecus* y *Dryopithecus* (que pertenecen a la línea evolutiva que condujo a los grandes antropomorfos africanos y a los humanos) en Grecia y en el oeste y centro de Europa, respectivamente. Estas representaciones se han confeccionado de manera parecida a como se realizan los retratos forenses.

del Mioceno

Durante el Mioceno, más de 100 especies de primates poblaron el Viejo Mundo. Nuevos fósiles indican que el antepasado de los grandes antropomorfos y de los humanos no procede de Africa, sino de Eurasia

David R. Begun

Reconstrucciones de los primates fósiles por John Gurche



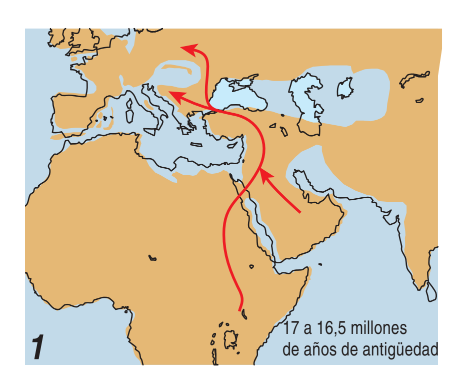
Ouranopithecus

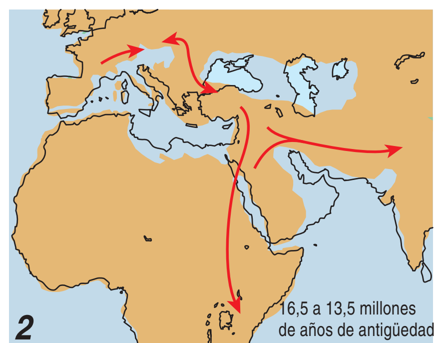


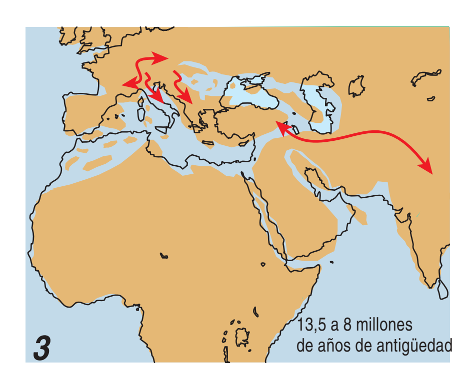
Oreopithecus

"Podemos, pues, afirmar con gran probabilidad que Africa fue habitada por especies que ya no existen, muy parecidas al gorila o al chimpancé; y como estas dos especies son las que más se asemejan al hombre, es un tanto más probable que nuestros antepasados viviesen en Africa en vez de en otro continente."

De esta forma reflexionaba Charles Darwin en su libro *El Origen del Hombre*, publicado en 1871. Aunque en aquel momento no se disponía de ningún fósil de simios antropomorfos o de humanos, los restos hallados desde entonces han confirmado ampliamente esta inteligente predicción sobre el origen de la humanidad. Sin embargo, la historia es más compleja de lo que Darwin imaginara. Los fósiles y los análisis genéticos indican que el último antepasado común de los humanos y de nuestro pariente más próximo, el chimpancé, vivió en Africa hace entre 6 y 8 millones de años. Pero, ¿de dónde procedía, a su vez, esta especie? Desde hace bastante tiempo, los paleoantropólogos han opinado que







LAURIE GRACE; FUENTE: F. RÖGL EN THE MIOCENE LAND MAMMALS OF EUROPE, DIRIGIDO POR G. RÖSSNER Y K. HEISSIG. VERLAG DR. FRIEDRICH PFEIL, 1999

también tendría raíces africanas. Sin embargo, cada vez existen más pruebas que indican que esta idea heredada es errónea.

En la actualidad el número de hominoideos es muy escaso (la superfamilia de los hominoideos agrupa a todos los simios antropomorfos, incluyendo gibones, siamang y los seres humanos). Pero hace entre 22 y 5,5 millones de años, durante el Mioceno, dominaron el mundo de los primates. Más de 100 especies distintas de hominoideos habitaban el Viejo Mundo, desde Francia a China en Eurasia y desde Kenia a Namibia en el continente africano. A partir de esta extraordinaria diversidad surgió el limitado número, en términos relativos, de los simios antropomorfos actuales. Sin embargo, sólo se han encontrado fósiles de los grandes antropomorfos (un grupo representado en la actualidad por chimpancés, gorilas y orangutanes; a los gibones y al siamang se los denomina pequeños antropomorfos) en Europa occidental y central, Grecia, Turquía, el sur de Asia y China. De esta forma, siguiendo la lógica de Darwin, Eurasia es el más probable lugar de nacimiento de la familia que agrupa a los grandes antropomorfos y a los humanos. (El término homínido se ha reservado tradicionalmente para los humanos y sus antepasados protohumanos, pero cada vez se utiliza más de forma que abarque también a los grandes antropomorfos, y se emplea "hominino" para referirse a humanos y protohumanos.)

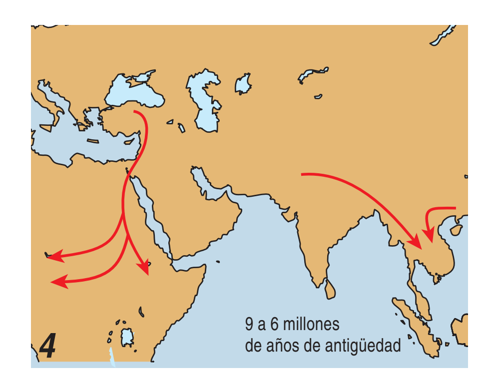
Quizá no debería sorprendernos que los hominoideos que dieron lugar a los homínidos evolucionasen en Eurasia y no en Africa. Los efectos combinados de migraciones, cambios climáticos, actividad tectónica y alteraciones ecológicas de una escala sin igual afectaron a esta zona desde el Mioceno y crearon un caldo de cultivo que fomentó los experimentos evolutivos de los hominoideos. El resultado fue toda una diversidad de hominoideos; dos de estas líneas evolutivas acabarían por encontrarse bien situadas para colonizar el sudeste asiático y Africa, y finalmente dar origen a los actuales grandes antropomorfos y a los humanos.

Largo trecho ha recorrido la paleoantropología desde que Georges Cuvier, el historiador natural francés que fundó la paleontología de vertebrados, escribiera en 1812 que "*l'homme fossile n'existe pas*" (el hombre fósil no existe). Extendía esta negación a todos los fósiles de primates. Aunque esa afirmación parece irrazonable hoy en día, en su época no existían muchas pruebas de que los primates hubieran vivido junto a otros animales extintos (como los mastodontes, los perezosos terrestres gigantes y los primitivos ungulados). Tiene su gracia que, algo más tarde, fuera el propio Cuvier quien describiese el que luego sería el primer fósil de primate que recibiría una denominación científica, *Adapis parisiensis* Cuvier 1822, un lémur hallado en una mina de caliza de París; lo confundió con un ungulado. En 1837, poco después de que falleciera, su discípulo Edouard Lartet reconoció y describió el primer fósil de primate antropoideo. Esta mandíbula —del primate hoy denominado *Pliopithecus*— encontrada en el sur de Francia y otros fósiles similares terminaron por convencer a los científicos de que animales de ese tipo habían habitado en el pasado los bosques primitivos de Europa. Unos 20 años más tarde, Lartet descubría en los Pirineos franceses el primer fósil de gran antropomorfo, *Dryopithecus*.

Durante los siguientes años del siglo XIX y hasta entrado el siglo XX, se recuperaron muchos más fragmentos de mandíbulas y dientes de

Resumen/*Revolución en los primates*

- Unicamente 5 géneros de simios antropomorfos sobreviven hoy en día. Se hallan confinados en pequeñas áreas de Africa y el sudeste asiático. Sin embargo, en el intervalo temporal transcurrido entre hace 22 millones y 5,5 millones de años vivieron en todo el Viejo Mundo docenas de géneros de primates.
- Desde hace mucho tiempo el antepasado de los grandes antropomorfos actuales y de los humanos evolucionó en Africa. Tal era la tesis recibida. Pero cada vez hay más pruebas que indican que, aunque Africa fue el lugar de aparición de los primeros hominoideos, el clado de los grandes antropomorfos y de los humanos se originó en Eurasia.
- A tenor del registro fósil los grandes antropomorfos actuales y los humanos descenderían de dos líneas evolutivas eurasiáticas: una, asiática, representada por *Sivapithecus* (probable antepasado del orangután), la otra, europea, por *Dryopithecus* (el antepasado más probable de los antropomorfos africanos y de los humanos).



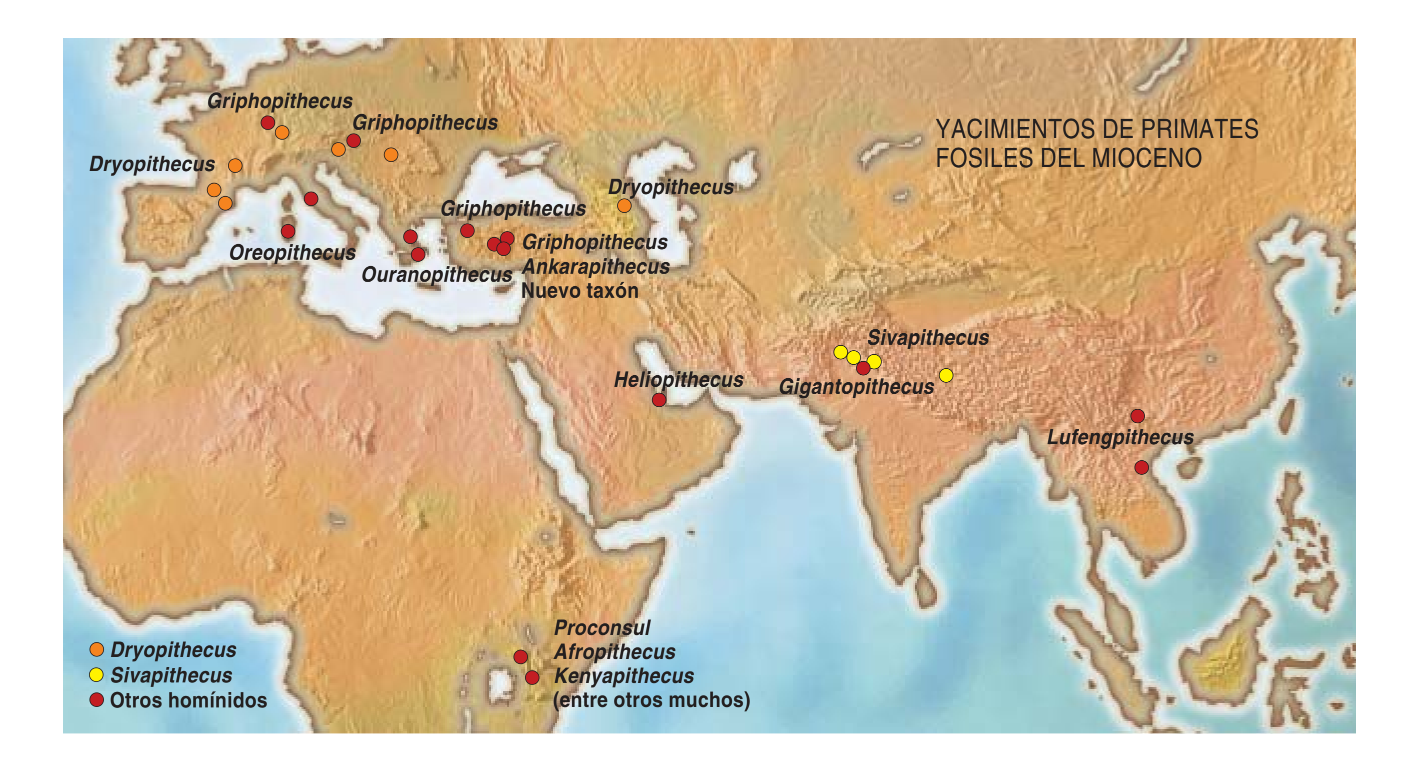
2. MIGRACION DE LOS PRIMATES. Africa fue la cuna del linaje primate. Allí aparecen los fósiles más antiguos, de hace 20 millones de años. Pero no mucho después se habían esparcido por todo el Viejo Mundo. Los cambios del nivel del mar unieron en algunas ocasiones Africa a Eurasia y en otras la aislaron. El papel de esos cambios en la evolución de los primates fue fundamental. Un puente de tierra tendido del este de Africa a Eurasia, hace entre 17 y 16,5 millones de años, permitió que los primates del Mioceno entrasen en Eurasia (*1*). Durante los siguientes millones de años se expandieron desde Europa occidental hasta el Lejano Oriente. Aparecieron los grandes antropomorfos; algunos de estos hominoideos primitivos regresaron a Africa (*2*). Aislados de Africa por la elevación del nivel de los mares, los grandes antropomorfos eurasiáticos evolucionaron en numerosas formas (*3*). Pero los cambios drásticos registrados en el clima, al final del Mioceno, extinguieron la mayoría de los hominoideos eurasiáticos. Las dos líneas evolutivas que sobrevivieron —representadas por *Sivapithecus* y *Dryopithecus*— lo consiguieron porque se desplazaron al sudeste asiático y al Africa tropical.

hominoideos, junto a algunos huesos de las extremidades, en España, Francia, Alemania, Austria, Eslovaquia, Hungría, Georgia y Turquía. Sin embargo, hacia la década de 1920, la atención pasó de Europa al Sur de Asia (India y Pakistán) y a Africa (principalmente a Kenia), debido a los espectaculares descubrimientos realizados en estas regiones. Los hominoideos de Eurasia cayeron en el olvido. Pero los hallazgos de fósiles a lo largo de los últimos veinte años han despertado nuevamente un gran interés por los fósiles de antropomorfos en Eurasia; los paleontólogos han recuperado especímenes tan completos como para que se pueda intentar reconstruir su aspecto y averiguar su parentesco con los primates y humanos actuales.

Los primates originales

Los investigadores han identificado más de 40 géneros de primates fósiles del Mioceno en yacimientos distribuidos por todo el Viejo Mundo (ocho veces el número de géneros que sobrevive en la actualidad). Esta variedad parece caracterizar a la familia de primates desde el principio: en cuanto aparecen en el registro fósil, su diversidad es grande; sólo del Mioceno inferior de Africa (entre los 22 y los 17 millones de años de antigüedad) se han reconocido 14 géneros. Y considerando que el registro fósil es incompleto, hay grandes probabilidades de que este número sea muy inferior al que realmente hubo en ese período.

Igual que los actuales, aquellos primates tenían tamaños muy diferentes. El más pequeño pesaba sólo tres kilogramos, poco más que un gato no muy grande; el mayor, con unos 80 kilos, alcanzaba el tamaño de los gorilas. En cuanto a su alimentación, la variedad era mayor que entre los primates actuales; algunos se alimentaban de hojas, otros de frutas y frutos secos, si bien la mayoría subsistía gracias a la fruta, como los hominoideos de hoy. La



¿A qué llamamos primate?

LOS PRIMATES ACTUALES —chimpancés, gorilas, orangutanes, gibones y siamang— y los seres humanos comparten un conjunto de caracteres que los diferencian del resto de los simios. Carecen de una cola visible; es mucho más importante de lo que parece porque significa que el torso y las extremidades deben reunir varios requisitos a fin de acometer ciertas tareas que antes efectuaba la cola. Los antropomorfos y los humanos tienen extremidades muy flexibles, que les permiten levantar los brazos por encima de su cabeza y colgarse. (Este el motivo por el que los antropomorfos tienen brazos mucho más largos y fuertes que las piernas; los seres humanos han modificado las proporciones entre sus miembros al volverse bípedos.) Por la misma razón, presentan un tórax ancho, la parte inferior de la espalda más corta, la pelvis y los tobillos muy móviles, los pies con gran capacidad prensil y una postura habitual más vertical que la de otros simios. Además, son más grandes, sobre todo los grandes antropomorfos (chimpancés, gorilas y orangutanes), los cuales, por otra parte, crecen y se reproducen con mayor lentitud. Los grandes antropomorfos y los humanos poseen los cerebros de mayor tamaño entre los primates y su inteligencia es mayor que la de cualquier mamífero en casi todas las maneras en que puede medirse (entre otras, el uso de utensilios, el reconocimiento de sí mismos ante un espejo, la complejidad social y las estrategias para conseguir alimentos).

Se considera primates fósiles a los que exhiban mayor parentesco con los antropomorfos actuales que con cualquier otro grupo. Entra dentro de lo esperado que las formas más antiguas no muestren tantas de las características que definen a los primates como las posteriores. *Proconsul*, procedente del Mioceno inferior, carece de cola, como demuestra la morfología de su sacro (situado en la base de la columna vertebral), donde se insertaría la cola si la hubiese tenido. Pero en *Proconsul* todavía no habían aparecido la movilidad de las extremidades, ni el tamaño relativo del cerebro de los hominoideos actuales. Pocos discrepan de que *Proconsul*, con 19 millones de años de antigüedad, sea el primer primate claro del registro fósil. La clasificación de otros "primates" del Mioceno inferior (*Limnopithecus*, *Rangwapithecus*, *Micropithecus*, *Kalepithecus* y *Nyanzapithecus*) es más problemática, en parte debido a la ausencia de restos diagnósticos del esqueleto poscraneal. Constituirían formas más primitivas, que vivieron antes de que se produjese la separación de las líneas evolutivas de los monos del Viejo Mundo y de los primates. En mi opinión, son primates por una razón principal: presentan en sus mandíbulas y dientes características similares a las de los primates.



MONO *PROCONSUL* GRAN ANTROPOMORFO

mayor diferencia entre los primeros hominoideos y los actuales estribaba en la postura y la forma en que se desplazaban. Los actuales exhiben un amplio repertorio de modos de locomoción, desde la acrobática braquiación de los gibones en los árboles hasta el terrestre "andar de nudillos" de los gorilas (más exactamente, apoyan el dorso de las segundas falanges); en cambio, los primeros hominoideos del Mioceno habían de desplazarse por las ramas de los árboles utilizando las cuatro extremidades.

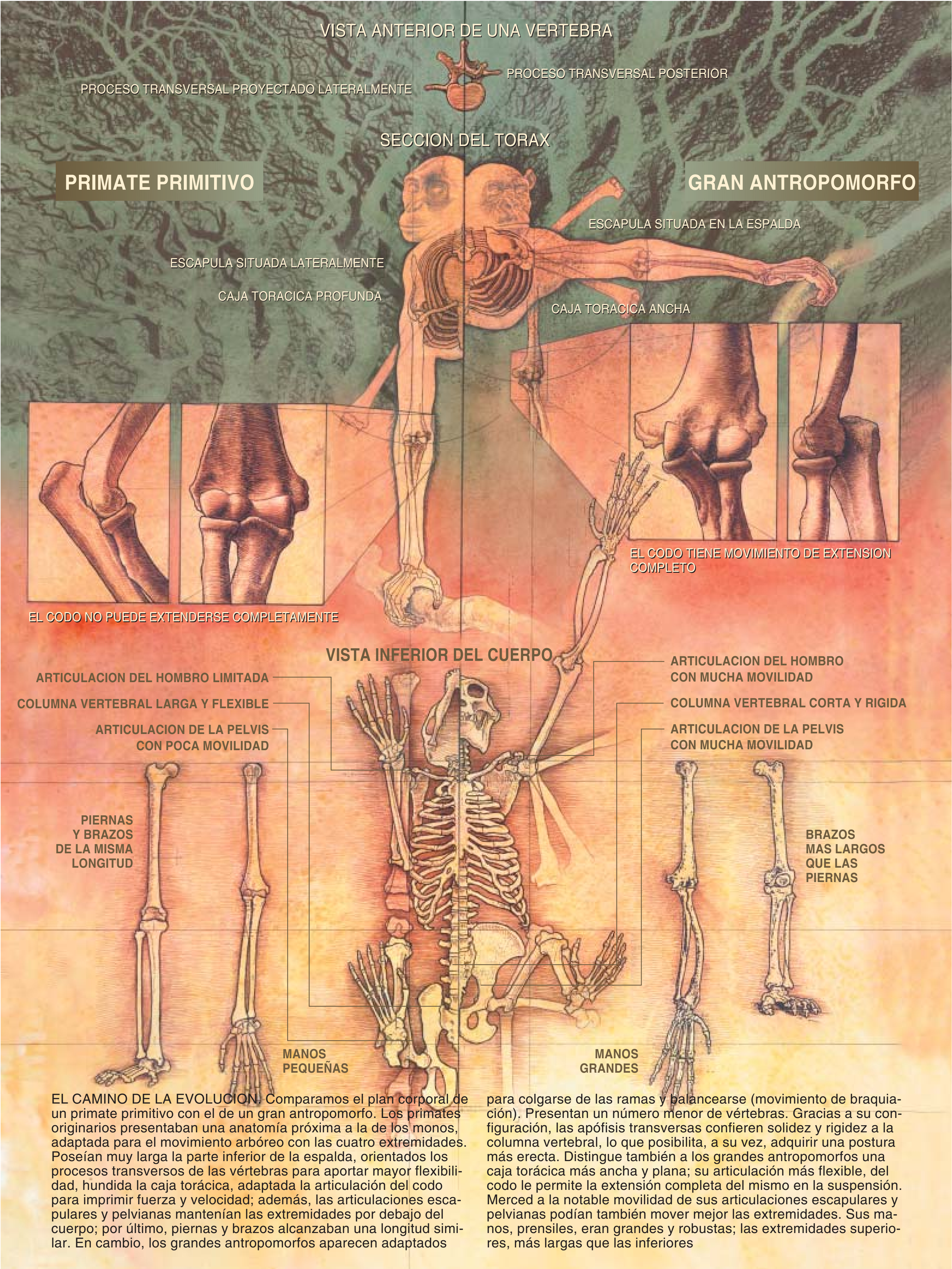
Para entender por qué la locomoción de los primeros hominoideos se restringía de esa manera, debemos tener en cuenta el plan corporal de los hominoideos del Mioceno temprano. La especie mejor conocida de ese período es *Proconsul*; se recuperaron fósiles excepcionalmente completos en los yacimientos de la isla Rusinga, en Kenia. En la actualidad, los especialistas reconocen cuatro especies de *Proconsul*, que abarcan pesos desde los 10 kilogramos hasta quizá más de 80 kilogramos. *Proconsul* puede proporcionar una buena imagen de la anatomía y locomoción de los primeros hominoideos. Al igual que los hominoideos modernos, carecía de cola. Comparado con los monos, tenía mayor movilidad en la pelvis, los hombros, codos, muñecas, manos y pies, presagiando las adaptaciones de los primates y los humanos de hoy que conceden una gran flexibilidad a esas articulaciones. En los primates modernos, esa movilidad mejorada permite un tipo de locomoción única: desplazarse de rama en rama colgándose de los brazos. En los humanos, la evolución ha puesto esta capacidad al servicio de la capacidad que tienen las extremidades superiores de manipular objetos. Tal exaptación de la movilidad hizo posible, entre otras cosas, que nuestros antepasados empezasen a fabricar utensilios.

Sin embargo, *Proconsul* y su grupo también conservan características primitivas, similares a las de los monos, en la columna vertebral, pelvis y extremidades anteriores. Estos rasgos les permitían, como a los simios que les antecedieron, desplazarse mejor sobre las ramas de los árboles que colgados o balanceándose en ellas. (Un género bastante enigmático de principios del Mioceno, encontrado en Uganda, *Morotopithecus*, plantea un problema interesante: parece que prefería suspenderse de los árboles, aunque las pruebas no son concluyentes.) Sólo cuando los primeros primates se libraron de esta herencia evolutiva pudieron ir adoptando las formas de locomoción de los primates contemporáneos.

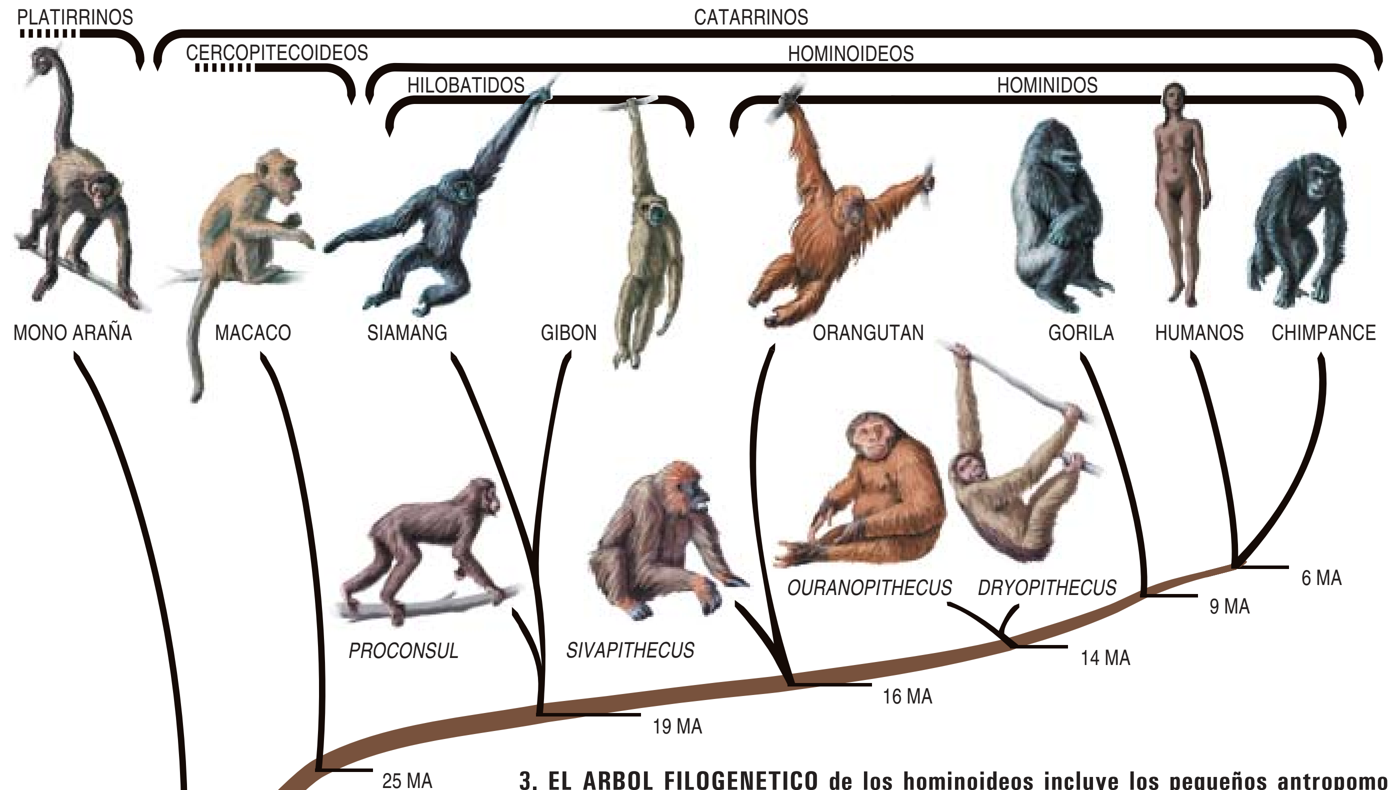
El paso a Eurasia

La mayoría de los primates del Mioceno temprano se extinguió. Pero uno de ellos, quizás *Afropithecus,* de Kenia, fue el antepasado de las especies que salieron por primera vez hacia Eurasia hace unos 16,5 millones de años. Alrededor de esa fecha, el nivel del mar de todo el planeta descendió, formándose un puente entre Africa y Eurasia. Se produjo un éxodo de mamíferos. Entre los que migraron fuera de su

PORTIA SLOAN (*esta página*); JOHN GURCHE (*página opuesta*)



EL CAMINO DE LA EVOLUCION. Comparamos el plan corporal de un primate primitivo con el de un gran antropomorfo. Los primates originarios presentaban una anatomía próxima a la de los monos, adaptada para el movimiento arbóreo con las cuatro extremidades. Poseían muy larga la parte inferior de la espalda, orientados los procesos transversos de las vértebras para aportar mayor flexibilidad, hundida la caja torácica, adaptada la articulación del codo para imprimir fuerza y velocidad; además, las articulaciones escapulares y pelvianas mantenían las extremidades por debajo del cuerpo; por último, piernas y brazos alcanzaban una longitud similar. En cambio, los grandes antropomorfos aparecen adaptados para colgarse de las ramas y balancearse (movimiento de braquiación). Presentan un número menor de vértebras. Gracias a su configuración, las apófisis transversas confieren solidez y rigidez a la columna vertebral, lo que posibilita, a su vez, adquirir una postura más erecta. Distingue también a los grandes antropomorfos una caja torácica más ancha y plana; su articulación más flexible, del codo le permite la extensión completa del mismo en la suspensión. Merced a la notable movilidad de sus articulaciones escapulares y pelvianas podían también mover mejor las extremidades. Sus manos, prensiles, eran grandes y robustas; las extremidades superiores, más largas que las inferiores



3. EL ARBOL FILOGENETICO de los hominoideos incluye los pequeños antropomorfos (siamang y gibones), los grandes antropomorfos (orangutanes, gorilas y chimpancés) y los humanos. La mayoría de los hominoideos del Mioceno corresponden a líneas evolutivas extintas, pero algunos cuentan entre los posibles antepasados de los antropomorfos actuales y de los humanos. Se cree que *Proconsul*, un primate primitivo del Mioceno, fue el último antepasado común de todos los hominoideos actuales; es muy probable que *Sivapithecus*, un gran antropomorfo primitivo, fuese el antepasado del orangután; y *Dryopithecus*, o bien *Ouranopithecus*, debió de dar lugar a la línea evolutiva que conduce a los antropomorfos africanos y a los humanos.

PORTIA SLOAN

Africa natal había elefantes, roedores, ungulados —como los cerdos y los antílopes— y algunos animales más exóticos, así aardvarks (cerdos hormigueros) y primates.

En el caso de los primates, de Africa pasaron a Eurasia a través de la península Arábiga, donde se han encontrado los restos de *Heliopithecus*, un primate similar a *Afropithecus*. Ambos, *Afropithecus* y *Heliopithecus*, a los que algunos científicos consideran miembros del mismo género, tienen en sus dientes una capa de esmalte gruesa, que les permite procesar alimentos duros, como los frutos secos, o de textura firme protegida por una cubierta resistente. Esta novedad en la morfología dentaria pudo contribuir mucho a que sus descendientes se establecieran en los bosques de Eurasia, ya que les permitía explotar una serie de recursos que no podrían haber aprovechado *Proconsul* ni la mayoría de los primeros hominoideos. Cuando medio millón de años más tarde el mar subió hasta cubrir el puente de tierra que unía Africa y Eurasia, los primates habían arraigado en sus nuevos territorios.

El desplazamiento de los organismos a nuevos ecosistemas conduce a la especiación. Así aconteció con la llegada de los primates a Eurasia. Más aún, en un abrir y cerrar de ojos a escala geológica, esos primates se adaptaron a nuevas condiciones ecológicas y se diversificaron en muchos géneros (conocemos al menos ocho en sólo 1,5 millones de años). Esta oleada de novedades evolutivas preparó el terreno para la aparición de los grandes antropomorfos y de los humanos. No hace mucho que se ha reconocido la importancia de Eurasia en esa eclosión. Se creía que primates más evolucionados que *Afropithecus* y *Heliopithecus* en su capacidad de procesar alimentos llegaron a Europa hace unos 15 millones de años, más o menos cuando aparecen en Africa, en concordancia con la hipótesis de que surgieron en ésta y posteriormente se dispersaron hacia el norte. En cambio, las nuevas pruebas fósiles apuntan a que otros primates más evolucionados (con mandíbulas muy robustas y grandes dientes) vivieron en Eurasia con bastante anterioridad a esa fecha. Otros compañeros y yo describimos en 2001 y 2003 un primate de morfología afín a los modernos, de 16,5 millones de años de antigüedad, hallado en yacimientos de Alemania y Turquía, *Griphopithecus*. Con ello, la presencia más antigua de primates en Eurasia retrocedía más de un millón de años.

El autor

DAVID R. BEGUN es profesor de antropología en la Universidad de Toronto. Obtuvo su doctorado en antropología física en la Universidad de Pennsylvania en 1987. Especializado en la evolución de los hominoideos durante el Mioceno, ha excavado en yacimientos de fósiles de España, Hungría, Turquía y Kenia.

El yeti

ALGUNAS PERSONAS, entre ellas científicos serios, han afirmado que la línea evolutiva del *Sivapithecus*, de la que desciende el orangután, tiene un representante vivo. Los detalles de la anatomía de ese animal cambian de un relato a otro, pero todos concuerdan en describirlo como un primate no humano enorme e hirsuto que camina erguido. Se dice que ha sido avistado en algunos lugares de Norteamérica y de Asia.

Por desgracia, hay más nombres de esta criatura que pruebas de su existencia: bigfoot, yeti, sasquatch, nyalmo, rimi, raksi-bombo, el abominable hombre de las nieves..., y la lista sigue creciendo.



SUPUESTAS PISADAS de pie de un yeti, fotografiadas en 1976 cerca de Coos Bay, Oregón (EE.UU.)

Aquellos que creen en la existencia del yeti (basándose en el hallazgo de pelos, heces, pisadas y algunas grabaciones de vídeo muy borrosas) a menudo apuntan al gran antropomorfo *Gigantopithecus* como su posible antepasado directo. *Gigantopithecus* fue probablemente dos o tres veces mayor que un gorila; se sabe que vivió hasta hace unos 300.000 años en China y el sudeste asiático.

No hay ninguna razón para que una criatura así no pudiese vivir hoy en día. Después de todo, gracias al registro subfósil sabemos que lémures del tamaño de gorilas medraron en la isla de Madagascar hasta que los humanos los abocaran a la extinción hace sólo mil años. El problema es que, aunque disponemos de fósiles de hace 20 millones de años correspondientes a primates del tamaño de un gato, no contamos ni con un solo hueso de este supuesto gran antropomorfo bípedo de casi media tonelada de peso, que viviría incluso en Norteamérica. Aunque todos los primatólogos y los paleontólogos de primates se alegrarían si el yeti fuese una criatura real, la ausencia absoluta de pruebas tangibles de su existencia hace que parezca de lo menos probable.

BETTMANN/CORBIS

En contra de la opinión mantenida a lo largo de mucho tiempo, según la cual de Africa procedieron todas las formas de primates, la aparente ausencia de esas especies más recientes en aquel continente entre los 17 y los 15 millones de años de antigüedad indica que algunos primates desarrollaron evolutivamente sus características craneales y dentales en Eurasia; luego volvieron a Africa, transformadas en especies más evolucionadas, tras un nuevo retroceso del mar. (Puede que algunos géneros —como *Kenyapithecus,* hallado en Fort Ternan, Kenia— desarrollaran algunas adaptaciones en su esqueleto poscraneal para vivir en el suelo, pero la mayoría de los primates guardaban un estrecho parecido, del cuello para abajo, con sus precursores del Mioceno inferior.)

Aparición de los grandes antropomorfos

A finales del Mioceno medio, hace unos 13 millones de años, había grandes antropomorfos en Eurasia. Destacan *Dryopithecus* en Europa (el gran antropomorfo fósil descrito por Lartet) y *Sivapithecus* en Asia. Al igual que los grandes antropomorfos actuales, estos animales tenían mandíbulas poderosas y largas, con grandes incisivos y caninos acuchillados (no cónicos), molares y premolares largos con superficies de masticación más bien simples. Un conjunto de características que indica la adaptación al consumo de frutos blandos y maduros. También tenían el mentón más corto, lo que refleja una menor importancia del olfato a favor de la visión. El análisis histológico de la dentadura de *Dryopithecus* y *Sivapithecus* indica que su crecimiento era bastante lento, como sucede en los grandes antropomorfos actuales. Probablemente, su vida seguía un curso similar al de los grandes antropomorfos, con una maduración lenta, vidas largas y una sola cría por parto. Según otros indicios, de existir hoy, su inteligencia no desentonaría con la de los grandes antropomorfos: la capacidad craneal de los fósiles de *Dryopithecus* es comparable con la de un chimpancé de un tamaño parecido. Aunque no tenemos un indicador directo para *Sivapithecus*, sabemos que el tamaño cerebral está muy relacionado con el tipo de desarrollo de los individuos a lo largo de su vida; por lo tanto, el de este primate no debió de diferir mucho de los tamaños actuales.

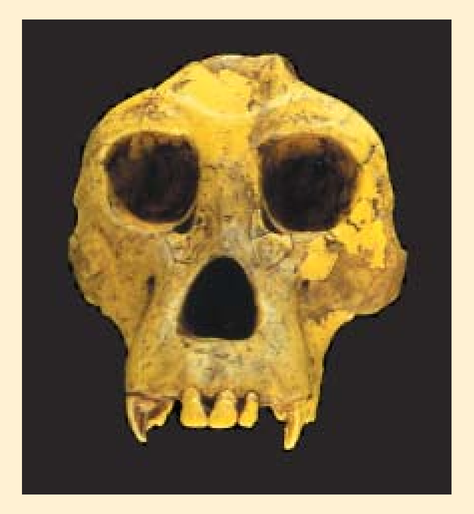
El estudio anatómico de las extremidades de ambos fósiles revela más características similares a las de los grandes antropomorfos. Merece destacarse una serie de adaptaciones a la suspensión de los árboles, sobre todo en la articulación del codo, que se extendía por completo y era estable a lo largo de todo su recorrido. Entre los simios, sólo los primates tienen este rasgo morfológico, al que deben en buena parte su capacidad de colgarse de las ramas y balancearse. Además, a los humanos nos proporciona la capacidad de lanzar objetos a gran velocidad y con precisión. *Dryopithecus* presenta otras muchas adaptaciones a la suspensión, en los huesos de las extremidades y en las manos y los pies, que le dotan de una gran capacidad de prensión. Estas características dan a entender que *Dryopithecus* se desplazaría por el dosel forestal de forma muy semejante a como lo hacen los grandes antropomorfos actuales. Conocemos menos el tipo

Golpes de suerte

A MENUDO, los descubrimientos de fósiles son el resultado de una combinación de pura suerte y de conjeturas bien fundadas. Este el caso de los hallazgos de dos de los fósiles de grandes primates más completos del registro fósil. Se dio con el primero de ellos en el yacimiento de Can Llobateres, en el término municipal de Sabadell. Este yacimiento ha proporcionado fragmentos de mandíbulas y dientes desde la década de 1940; a finales de los años ochenta fui invitado por unos paleontólogos españoles a reanudar las excavaciones en aquel lugar. Volví para una segunda campaña de excavaciones al año siguiente.

Mientras mis colegas y yo estábamos limpiando el sedimento superficial apareció un premolar de hominoideo. Vimos asombrados cómo rodó por la pendiente, a cámara lenta nos pareció, hasta detenerse a nuestros pies. Pocos días después habíamos recuperado la primera cara, casi completa, de *Dryopithecus* (*véase la ilustración*), el fósil de antropomorfo más completo de Can Llobateres en el medio siglo que se lleva excavando allí. Luego escudriñamos el mismo nivel de sedimentos alrededor del yacimiento y encontramos algunos fragmentos de extremidades en otra área; tras una excavación más exhaustiva el año siguiente, constituirían el esqueleto más completo de *Dryopithecus* descubierto hasta la fecha.

Nueve años después, en Hungría, iniciamos una nueva excavación en el yacimiento de Rudabánya, que había venido proporcionando numerosos fósiles de *Dryopithecus* —sobre todo dientes y restos de esqueleto—. Debíamos concentrar nuestros esfuerzos en una capa de sedimentos oscuros que indicaban un mayor contenido en materia orgánica, característica que a menudo anuncia riqueza fosilífera. Este nivel era visible en la sec-




FOSILES EXCEPCIONALES de *Dryopithecus*, uno de los primeros grandes antropomorfos, procedentes de yacimientos de España (*a la izquierda*) y de Hungría (*a la derecha*).

DAVID R. BEGUN

de locomoción de *Sivapithecus*. Algunas características de la anatomía de sus extremidades sugieren que se suspendía de los árboles; en cambio, otras indican que tendría hábitos cuadrúpedos. Es casi seguro que se valía de un modo de locomoción del que no tenemos hoy un análogo, derivado de unas circunstancias ecológicas singulares.

Descendencia de *Sivapithecus* y *Dryopithecus*

La línea evolutiva de *Sivapithecus* medró en Asia, con descendientes en Turquía, Pakistán, India, Nepal, China y el Sudeste asiático. La mayoría de los análisis filogenéticos coinciden en que el orangután actual (*Pongo pygmaeus*) desciende de *Sivapithecus*. Este gran antropomorfo, que habita las selvas húmedas de Borneo y Sumatra, es el único representante de un grupo que tanto llegó a extenderse.

También fue espléndida la radiación de los grandes antropomorfos en el oeste. Tras la primera especie de *Dryopithecus*, *D. fontani* (la descubierta por Lartet), aparecieron varias en los siguientes tres millones de años. Algunos de sus descendientes tuvieron características más especializadas. En dos millones de años, evolucionaron cuatro nuevas especies de *Dryopithecus*. Se expandieron desde el noroeste de España hasta la República de Georgia. Sin embargo, se sigue debatiendo la posición de *Dryopithecus* en el árbol filogenético de los hominoideos. Algunos estudios lo relacionan con los grandes antropomorfos asiáticos; para otros sería el antepasado de todos los grandes antropomorfos aún existentes. El análisis filogenético que he realizado de estas especies, el más completo en cuanto al número de características anatómicas consideradas, indica que *Dryopithecus* revela un estrecho parentesco con *Ouranopithecus*, un primate procedente de Grecia. Uno de estos dos géneros europeos sería el antepasado de los antropomorfos africanos y de los humanos.

En 1999 descubrimos un cráneo de *Dryopithecus* en Rudabánya (Hungría) que refuerza esta idea. Apodado "Gabi" en honor a su descubridor, el geólogo húngaro Gabor Hernyák, es el primer fósil que conserva un elemento anatómico esencial: la conexión entre la cara y la caja craneana, o neurocráneo. Gabi muestra que *Dryopithecus* tenía un neurocráneo largo y bajo, como el de los antropomorfos africanos y los humanos fósiles; también era su región nasal plana y la parte baja de la cara grande. Y quizá lo que reviste mayor interés: al igual que los antropomorfos africanos y los primeros humanos, *Dryopithecus* presentaba clinorrinquia (vista de lado, la cara se orienta hacia abajo). En cambio, los orangutanes, *Proconsul*, los gibones y el siamang muestran la cara orientada hacia arriba, una disposición conocida como aerorrinquia. Este aspecto fundamental de la arquitectura craneal de *Dryopithecus* nos habla a favor de una relación evolutiva estrecha

ción norte-sur del yacimiento; se volvía más claro hacia el norte, donde, según mi opinión, era menos probable que contuviese fósiles. Pedí a Gabor Hernyák, un geólogo húngaro, que empezase por el extremo norte y avanzase hacia el sur, donde supuestamente se encontraría el filón de huesos. Pero menos de un minuto más tarde, Gabor me llamó excitado. Quería que volviese al lugar donde le había dejado. Allí, en lo que parecía un sedimento pobre en contenido fósil, había descubierto un pequeño fragmento de un maxilar de *Dryopithecus*. Cuando finalizamos la extracción del fósil, teníamos ante nosotros el cráneo de *Dryopithecus* más completo jamás descubierto, el primero donde todavía estaba la cara unida al cráneo (*véase la ilustración*).

Este cráneo de Rudabánya —apodado "Gabi" en honor a su descubridor— ilustra mejor que ningún otro fósil el estrecho parentesco que hay entre *Dryopithecus* y los grandes antropomorfos africanos. Siempre recordaré la mirada de mi amigo, y codirector de las excavaciones, László Kordos cuando volvimos a la ciudad. Estaba escribiendo un correo electrónico y se giró hacia mí con cara aburrida. Me preguntó si había algo nuevo. "Poca cosa", le contesté. "Sólo hemos encontrado un cráneo de *Dryopithecus*".

con la línea de los antropomorfos africanos y los humanos. Un indicio más de esta relación es el parecido del cráneo de *Dryopithecus* con los cráneos infantiles o juveniles de los chimpancés, coincidencia común en las morfologías ancestrales. El aspecto característico de los cráneos adultos de chimpancés, gorilas y humanos fósiles evolucionó a partir del plan básico que representan *Dryopithecus* y los juveniles de los antropomorfos africanos.

Oreopithecus

Otro género de primate miocénico merece especial atención. El primate eurasiático mejor conocido, en cuanto al porcentaje de esqueleto recuperado, es el *Oreopithecus,* de 7 millones de años de antigüedad hallado en Italia. Fue descrito en 1872 por Paul Gervais. *Oreopithecus* estaba más especializado en el consumo de hojas que cualquier otro fósil, mono o primate, del Viejo Mundo. Sobrevivió hasta muy avanzado el Mioceno en los bosques, densos y aislados, de las islas de la Toscana, que terminaron por unirse entre sí y con el resto de Europa, cuando el mar retrocedió, para formar la columna vertebral de la península itálica. Esta especie tenía un cuerpo de gran tamaño y un cerebro pequeño, pero su anatomía es tan extraña que no queda claro si se trata de una forma primitiva que precede a la divergencia de los gibones y los grandes antropomorfos, si es un gran antropomorfo temprano o un pariente cercano de *Dryopithecus*. Según Meike Köhler y Salvador Moyà Solà, del Instituto Paleontológico Miquel Crusafont de Sabadell, *Oreopithecus* podía caminar de forma bípeda y tenía una mano similar a la humana, prensil. Sin embargo, la mayoría de los paleoantropólogos opina que, muy al contrario, se suspendía de los árboles. Fuera como fuese, *Oreopithecus* testimonia la gran diversidad y éxito evolutivo en su adaptación a nuevos entornos de estos primates eurasiáticos.

Pero, ¿qué fue de la multitud de especies que no evolucionó en la senda de los grandes antropomorfos actuales y los seres humanos? ¿Por qué perduraron los antepasados de las especies actuales? Parte de la respuesta nos la da la climatología. Durante el Mioceno medio, los grandes antropomorfos prosperaron en Eurasia gracias a la amplia cubierta de unas selvas subtropicales y a unas temperaturas siempre cálidas. Estas condiciones aseguraban la producción constante de frutos carnosos y un hábitat arbóreo muy diverso, con varios estratos.

Los cambios climáticos del Mioceno superior pusieron fin a esta vida regalada. El efecto combinado de la formación de los Alpes, el Himalaya y las montañas de Africa oriental, la modificación de las corrientes oceánicas y el inicio de la aparición de los casquetes helados en los polos provocaron el nacimiento de los ciclos monzónicos en Asia, la desecación del este de Africa y unas temperaturas más templadas en Europa. El cambio general en los ecosistemas fue la causa de la extinción de los grandes primates eurasiáticos. Las dos líneas evolutivas que sobrevivieron —las representadas por *Sivapithecus* y *Dryopithecus*— lo lograron gracias a que se desplazaron al sur del Trópico de Cáncer: hacia el sudeste de Asia desde China, hacia el Africa tropical desde Europa. Buscaban las condiciones ecológicas a las que estaban tan bien adaptados en Eurasia.

El modelo biogeográfico antes esbozado proporciona también una perspectiva muy importante sobre una cuestión largamente debatida en paleoantropología, el cómo y el porqué de que los humanos adquiriésemos la locomoción bípeda. Para resolver este problema hemos de conocer a partir de qué tipo de locomoción se desarrolló la bipedestación. Pero carecemos de pruebas fósiles claras del primer bípedo y de su antepasado. Se ignora el tipo de locomoción ancestral. Se han propuesto, sin embargo, dos hipótesis: la locomoción bípeda apareció a partir de animales que trepaban por las ramas o se colgaban de ellas; surgió de una forma de locomoción terrestre, quizás a partir del "andar apoyados en los nudillos".

Nuestros ancestros entre los grandes antropomorfos

El antepasado eurasiático de los antropomorfos africanos y de los humanos emigró hacia el sur ante la aridez mayor y el enfriamiento de sus ecosistemas, que causaron el reemplazo de los bosques cerrados por otros abiertos y por praderas. Las adaptaciones a la vida en el suelo de este linaje (en particular al "andar de nudillos") desempeñaron un papel fundamental en que resistiese la pérdida del hábitat forestal y pudiera regresar a Africa. Una vez allí, algunos primates volvieron a la selva, otros se establecieron en ambientes arbolados variados, y uno de ellos, el antepasado de los humanos, al adoptar una vida que transcurría sólo en el suelo, se aventuró en campo abierto.

La flexibilidad adaptativa se manifiesta una vez y otra en la evolución de los primates y de los seres humanos. Los antropomorfos del

Mioceno inferior abandonaron Africa pertrechados de nuevas adaptaciones de sus mandíbulas y dientes que les permitieron explotar nuevos recursos. Gracias a toda una serie de adaptaciones esqueléticas, pudieron los grandes antropomorfos eurasiáticos vivir en ecosistemas muy diversos. Además, su notable cerebro les facultaba para abordar complejos problemas sociales y ecológicos. Estas modificaciones hicieron posible que algunos sobrevivieran a los profundos cambios climáticos del final del Mioceno y regresasen a Africa, hace unos nueve millones de años. La línea evolutiva que dio lugar a los antropomorfos africanos y a los humanos estaba preadaptada, pues, para superar las dificultades de un ecosistema que cambiaba de manera radical. No es sorprendente que una de estas especies desarrollara un gran cerebro y una técnica muy elaborada.

Hace más de 20 años, mientras estudiaba la carrera, empecé a prestar mi atención a los primates fósiles convencido de que sólo entenderíamos por qué habían aparecido los humanos si llegábamos a saber a partir de qué, cuándo, dónde y cómo evolucionaron. Por lo usual, es en los antropomorfos actuales donde se busca una aproximación a la anatomía y comportamiento de los primeros humanos. Así hemos aprendido mucho, pero los grandes antropomorfos actuales también han evolucionado desde unos orígenes. El estudio de los primates fósiles nos proporciona una perspectiva única de los antepasados de los grandes antropomorfos y de los humanos y una forma de conocer el punto de inicio de los procesos y circunstancias que llevaron a la existencia de este grupo. Así, al haber establecido la conexión entre los grandes primates fósiles eurasiáticos y los grandes antropomorfos africanos actuales y los humanos, podemos reconstruir el último antepasado común de los chimpancés y los humanos: fue un primate, parecido al chimpancé, que vivía en la selva, "andaba de nudillos", comía frutas, empleaba útiles, cazaba animales y vivía, como los chimpancés y los seres humanos, en grupos sociales muy dinámicos y complejos.

Ramificaciones tortuosas

Todavía nos queda mucho por conocer. Numerosos primates fósiles sólo están representados por mandíbulas y dientes; tenemos poca o ninguna información sobre su postura habitual y forma de locomoción, su tamaño cerebral o corporal. No se han encontrado todavía restos fósiles de los antepasados de los antropomorfos africanos. Y existe un gran vacío geográfico y cronológico en el registro fósil entre los representantes de los primeros miembros del linaje homínido africano en Europa (*Dryopithecus* y *Ouranopithecus*) y los primeros homínidos fósiles africanos.

Si nos desplazamos más arriba por nuestro árbol familiar (que quizá deberíamos llamar mejor arbusto familiar), puede que experimentemos cierta confusión: no nos parecerán muy humanos los presuntos primeros miembros de la familia humana. Por ejemplo, se ha descubierto hace poco en el Chad un fósil de 6 o 7 millones de años de antigüedad denominado *Sahelanthropus tchadensis*. Tiene caninos pequeños, como los humanos y, quizá, más centrado el foramen magnum (el agujero en la base del cráneo que conecta el cerebro con la médula espinal), lo que podría indicar que esta especie era bípeda. Pero *Sahelanthropus* muestra muchas características similares al chimpancé: un cerebro de pequeño tamaño, una cara proyectada hacia delante, la parte posterior del cráneo orientada hacia atrás y músculos occipitales muy desarrollados.

Otros fósiles de hace 6 millones de años, descubiertos en Kenia y asignados a la especie *Orrorin tugenensis,* ofrecen también un parecido mosaico de rasgos de chimpancé y humano; lo mismo se ve en *Ardipithecus ramidus kadabba,* fósil de 5,8 millones de años de antigüedad hallado en Etiopía. Cada uno de estos taxones ha sido considerado por sus descubridores como un antepasado de los humanos. Pero realmente no sabemos lo suficiente de estos taxones para afirmar si son protohumanos, antepasados de los antropomorfos africanos o meras formas terminales extintas. En mi opinión, el primer fósil que sin duda pertenece al grupo de los humanos es *Ardipithecus ramidus ramidus*, de 4,4 millones de años, desenterrado en Etiopía.

La idea de que los antepasados de los grandes antropomorfos y de los humanos aparecieron en Eurasia se halla sujeta a controversia, pero no porque carezca del apoyo fósil necesario. El legado de Darwin inspira ese escepticismo. Según la interpretación habitual de su predicción reproducida al principio de este artículo, la evolución de humanos y antropomorfos africanos debería haber ocurrido sólo en Africa. Otros dudan porque se atienen al aforismo que reza: "La ausencia de prueba no es prueba de la ausencia". O sea, que no hayamos encontrado fósiles de grandes antropomorfos en Africa no quiere decir que no existan. Esto es cierto. Pero hay numerosos yacimientos paleontológicos en Africa de entre 14 y 7 millones de años de antigüedad que han proporcionado abundantes restos de animales de hábitat forestal; ninguno contenía fósiles de grandes antropomorfos. Aunque es posible que los primates eurasiáticos evolucionaran en paralelo con un antepasado africano aún no descubierto, tal fenómeno parece improbable.

Para reconstruir la historia evolutiva de nuestra extensa familia sería de gran ayuda disponer de un registro fósil más completo. Y los trabajos de excavación que están en marcha son muy prometedores y llenarán algunos de los vacíos de nuestro conocimiento. Pero hasta entonces debemos elaborar hipótesis con lo que ya sabemos.

Bibliografía complementaria

FUNCTION, PHYLOGENY AND FOSSILS: MIOCENE HOMINOID EVOLUTION AND ADAPTATIONS. Dirigido por David R. Begun, Carol W. Ward y Michael D. Rose. Plenum Press, 1997.

THE PRIMATE FOSSIL RECORD. Dirigido por Walter Carl Hartwig. Cambridge University Press, 2002.

RUDABÁNYA: A LATE MIOCENE SUBTROPICAL SWAMP DEPOSIT WITH EVIDENCE OF THE ORIGIN OF THE AFRICAN APES AND HUMANS. László Kordos y David R. Begun en *Evolutionary Anthropology*, vol. 11, n.º 2, págs. 45-57, 2002.

Las estrellas binarias

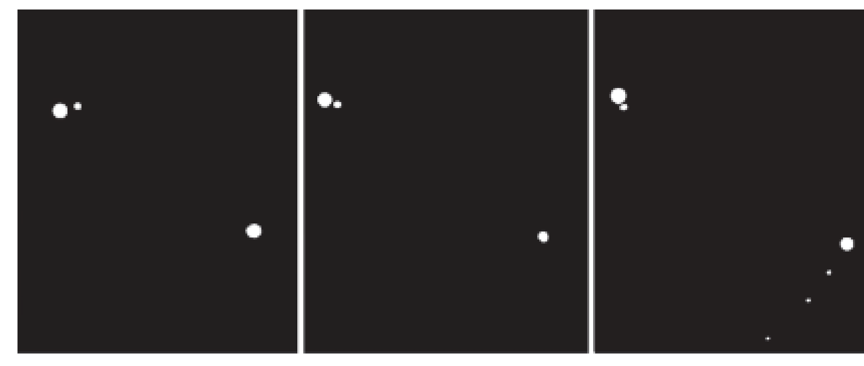
Al descubrir su verdadera naturaleza, William Herschel extendió las leyes de la física más allá del sistema solar. Hoy día son un verdadero laboratorio de astrofísica

Teodoro Vives

El jesuita Giovanni Baptista Riccioli observaba con un telescopio en 1650, en Palermo, que la estrella Mizar de la Osa Mayor aparecía doble, con una separación angular de 14" entre las dos componentes. Una de ellas se revelaría a su vez como un sistema de dos estrellas casi dos siglos y medio más tarde, en 1889, gracias a las observaciones espectroscópicas del norteamericano Edward C. Pickering. Sería la primera binaria espectroscópica conocida, es decir, descubierta por el análisis de su espectro y no por medio de una observación directa. Entre tanto, fueron apareciendo más estrellas dobles o múltiples. Christian Huyghens había visto ya en 1656 que θ Orionis, en el Trapecio de Orión, se resolvía en varias estrellas; en 1664 Robert Hooke comprobó que el telescopio desdoblaba γ Arietis. Durante el mismo siglo XVII, dos misioneros jesuitas hallaron un par de estrellas dobles más: α Crucis, descubierta por el padre Fontenay en el Cabo de Buena Esperanza en 1685, y α Centauri, descubierta por el padre Richaud en 1689 cuando observaba un cometa en Pondicherry, en la India.

Para 1781, Christian Mayer había descubierto con su telescopio la duplicidad de hasta 89 estrellas y aventurado la hipótesis de que algunas pudieran ser sistemas físicos reales. Sin embargo, la opinión general era que se trataba de un simple fenómeno de perspectiva óptica: las dos estrellas, separadas en realidad por una gran distancia, aparecían por azar muy próximas en la visual. Las dos componentes de una binaria suelen ser estrellas de magnitudes aparentes muy distintas, una brillante y la otra débil; puesto que se creía que la luminosidad intrínseca de todas las estrellas tenía que ser más o menos igual, se pensaba que una estaría cerca y la otra lejos. Por eso, aunque John Michell había demostrado en 1767, basándose en el cálculo de probabilidades, que la existencia de un número tan grande de estrellas dobles no podía deberse al azar, las cábalas de Mayer sobre la posibilidad de que una estrella pequeña girase en torno de otra mayor no fueron tomadas en serio por casi nadie.

En 1779 William Herschel, que al principio tampoco creía en la existencia física de las binarias, observó esos pares de estrellas que parecían estar juntas en el cielo con la finalidad de determinar paralajes. La idea de Herschel era la siguiente: si, de las dos estrellas que aparecen juntas, la más brillante está próxima y la más débil muy lejos, será posible medir el desplazamiento paraláctico de la estrella próxima respecto a la más distante, originado por el movimiento de la Tierra alrededor del Sol a lo largo del año; y, por tanto, determinar trigonométricamente su distancia. Las observaciones dieron un resultado muy diferente del esperado. Herschel no consiguió observar el desplazamiento paraláctico anual, pero sí un movimiento de ro-



1. EN LA PARTE SUPERIOR de estas tres fotografías aparece la estrella binaria visual Krüger 60. Dista 13 años luz de nosotros. Se observa el giro de sus dos componentes respecto a la estrella fija en la parte inferior derecha. Las fotos se han tomado a lo largo de 12 años.

OBSERVATORIO DE YERKES

tación de una componente alrededor de la otra. Durante 25 años Herschel observó pacientemente más de 50 estrellas dobles. Llegó a la conclusión de que se trataba realmente de sistemas binarios ligados por la gravedad.

En algunos casos fue posible obtener los períodos orbitales de revolución. Para Cástor, por ejemplo, Herschel encontró un período de 342 años, bastante menor que los 467 admitidos hoy. Cástor es una estrella brillante de la constelación de Los Gemelos, bien visible desde nuestras latitudes durante las noches de invierno. A simple vista se presenta como un punto de luz de 1,5 magnitudes estelares aparentes. Con un pequeño telescopio de aficionado se desdobla en dos refulgentes estrellas, Cástor A y Cástor B, de 2,0 y 2,8 magnitudes respectivamente, separadas unos tres segundos de arco. Son las componentes de la binaria que observó Herschel con su telescopio. Lo que no podía sospechar es que cada una de esas estrellas era a su vez una binaria; así Cástor A y Cástor B forman un sistema de cuatro estrellas unidas entre sí por la gravedad. Hubo que esperar hasta 1905 a que Hebert D. Curtis, del Observatorio Lick, probase su carácter binario a partir de las oscilaciones de las rayas espectrales de ambas componentes. Pero las sorpresas de Cástor no acabaron ahí.

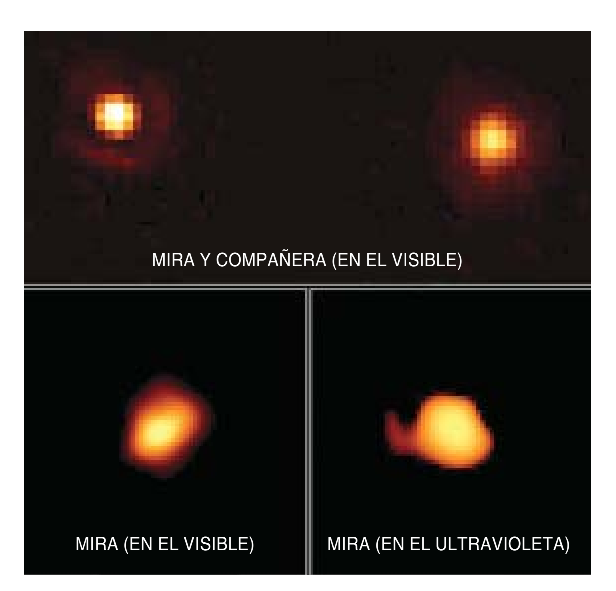
Cerca de Cástor, a unos 70 segundos de arco, se observa una débil estrellita de 9ª magnitud cuya luminosidad varía con un período de casi 20 horas. En 1916 W. S. Adams y A. H. Joy, del Observatorio de Monte Wilson, determinaron que se trataba de una binaria espectroscópica; en 1926 H. Van Gant, del Observatorio de Leiden, descubría que era una binaria eclipsante; es decir, que las variaciones luminosas se debían a que las dos componentes se eclipsaban periódicamente.

Pero lo más notable es que las observaciones han mostrado que esta estrella variable (YY Geminorum) está unida por la gravedad a Cástor; ha recibido el nombre de Cástor C. Cástor constituye, en realidad, un sistema de seis estrellas; gravitan unas alrededor de otras y todas en torno a un centro de masas común.

Con el descubrimiento de sistemas binarios compuestos por dos estrellas que giraban en torno a su centro de masas, Herschel demostraba en 1803 que la ley física de la gravitación de Newton se podía aplicar más allá del sistema solar. Fue en realidad el nacimiento de la astrofísica moderna.

Tipos de binarias

El último catálogo de Jeffers, el de 1963, contiene 64.246 estrellas binarias. Sólo se han estudiado las órbitas de unas 600, debido a que las observaciones han de cubrir por lo menos la mitad del período para proporcionar elementos orbitales bien determinados.



2. LA GIGANTE ROJA MIRA, u Omicron Ceti (*arriba, a la derecha*), tiene una compañera enana blanca (*arriba, a la izquierda*); el Hubble logró esta resolución del sistema binario en 1996. De Mira sale una especie de gancho, en dirección hacia la compañera (*abajo, a la derecha*): quizá sea material atraído hacia ésta, quizá material de la atmósfera de la gigante calentado por la enana. La elongación asimétrica de Mira en el visible *(abajo, a la izquierda)* puede deberse a deformaciones engendradas por el ciclo de expansiones y contracciones que le confiere su carácter de estrella variable o a erupciones no resueltas en la imagen.

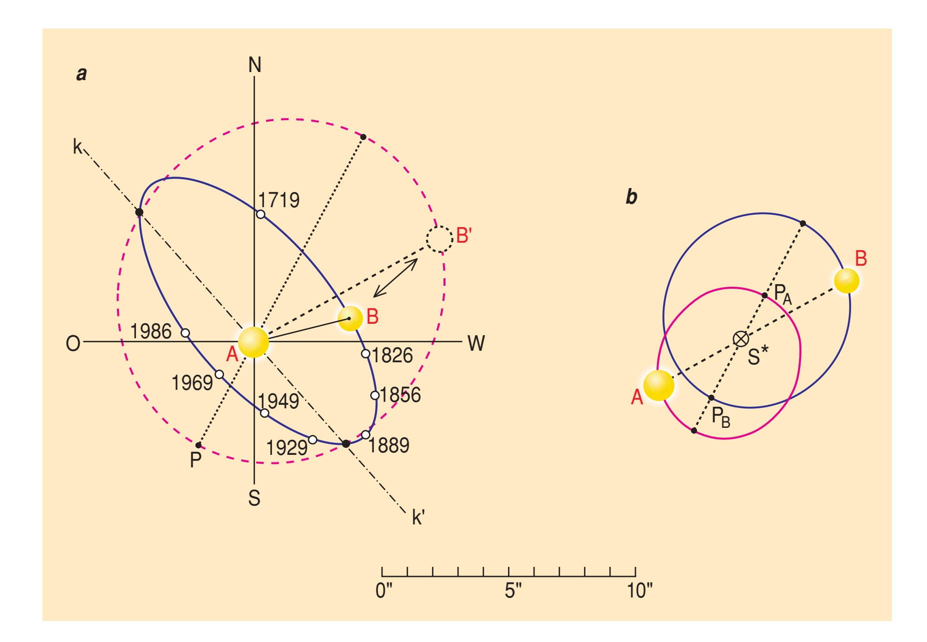
M. KAROVSKA Y NASA, ST ScI-PRC97-26

Pero el resultado más sorprendente del estudio de las estrellas binarias es que no constituyen casos excepcionales. Muy al contrario, la mayoría de las estrellas pertenecen a sistemas binarios e incluso múltiples. De las estrellas cercanas al Sol, hasta una distancia de 5 parsec, casi un 60 % pertenecen a un sistema doble. Y hasta una distancia de 25 parsec, un 44 %. Cuanto más lejana se halla una estrella doble, tanto más difícil resulta descubrir su carácter binario. Hoy día, teniendo en cuenta este efecto de selección observacional, se admite que la proporción de un 60 a 70 % continúa siendo válida en grandes regiones de nuestra Galaxia.

Por otra parte, existen tipos especiales de estrellas que con toda probabilidad son siempre binarias. Es el caso, por ejemplo, de las estrellas Wolf-Rayet, las estrellas de tipo espectral A con líneas metálicas (Am) o con espectro peculiar (Ap), las novas, las variables U Geminorum y, quizá, las estrellas simbióticas de espectro anómalo.

La naturaleza binaria de los sistemas distantes se puede descubrir mediante el corrimiento Doppler de las líneas espectrales (binarias espectroscópicas) y, en ciertos casos fotométricamente: cuando el plano orbital pasa más o menos por la posición del observador terrestre de manera que las dos estrellas se eclipsen de forma sucesiva y originen mínimos en la luminosidad total de las dos estrellas. Por eso, de acuerdo con las técnicas de observación utilizadas en cada caso, las estrellas binarias suelen clasificarse en binarias visuales (a través del telescopio se ven las dos estrellas y se observa el movimiento relativo de una alrededor de la otra), espectroscópicas y eclipsantes.

Cuando las dos estrellas de un sistema binario están tan alejadas entre sí que media sólo la interacción gravitatoria, se tiene una binaria desacoplada. Pero si se encuentran muy próximas, las fuerzas de marea las distorsionarán y lle-

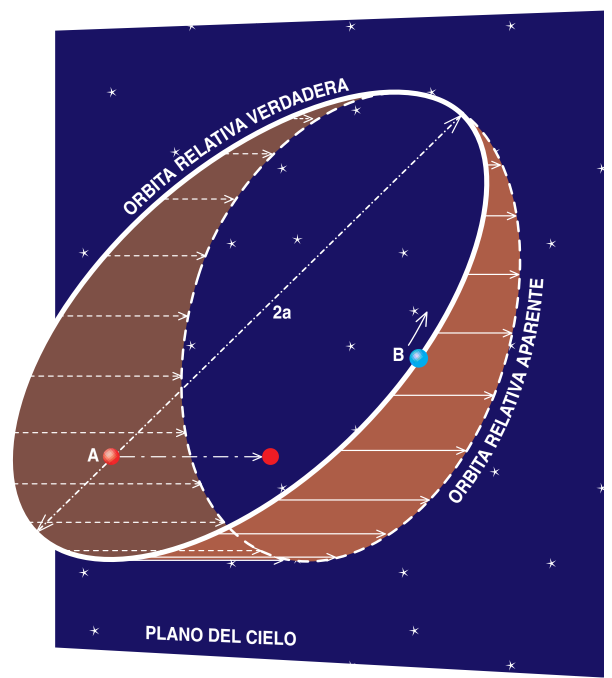


3. ORBITAS KEPLERIANAS aparentes y verdaderas de Cástor A y Cástor B. La elipse de trazo continuo es la órbita aparente descrita por Cástor B (B) alrededor de Cástor A (A), trazada con observaciones astronómicas desde 1719 hasta 1986; la elipse discontinua es la órbita relativa verdadera calculada a partir de la aparente (*a*). Ambas elipses se hallan en planos diferentes, pero tienen en común la línea de los nodos kk'. Las órbitas absolutas de las dos estrellas en torno al centro de masas S* se ven en *b*. Las dos figuras se han dibujado a la misma escala.

garán a arrancar gas de una para arrojarlo a la otra. Se tratará entonces de una binaria semidesacoplada. Si llegan prácticamente a tocarse hablamos de binaria de contacto. En el caso de que una de las componentes de un sistema semidesacoplado o incluso de contacto constituya una enana blanca, una estrella de neutrones o un agujero negro, la materia que cae a gran velocidad sobre la misma producirá una radiación muy energética (rayos X o gamma).

Si consideramos una estrella aislada en el espacio, lejos de las demás estrellas o de otros cuerpos masivos, nos resultará casi imposible averiguar directamente su masa. De ahí la importancia de los sistemas estelares dobles, las estrellas binarias, que obedecen las leyes de la gravedad y proporcionan el único medio directo de determinar masas estelares. Tras observar las estrellas binarias durante años, se ha establecido una relación masa-luminosidad que permite estimar indirectamente la masa de estrellas no binarias.

El movimiento de las dos estrellas de un sistema binario sigue las leyes de Kepler. En el caso de las binarias visuales, se obtiene la masa del sistema aplicando esas leyes a las observaciones. Cada una de las dos estrellas describe una órbita absoluta en torno al centro de masas. Estas órbitas, elipses, obedecen a la ley de las áreas de Kepler en un sistema de referencia inercial. La trayectoria relativa descrita por una estrella en torno a la otra dibuja también una elipse kepleriana, llamada órbita relativa. Un observador en la Tierra ve la órbita verdadera, absoluta o relativa, proyectada sobre un plano perpendicular a la visual; la determinación de los elementos orbitales consiste en pasar de esa órbita aparente observada a la órbita real.



4. SOBRE EL PLANO PERPENDICULAR a la visual del observador terrestre, el plano del cielo, se proyecta la órbita relativa aparente que observan los astrónomos. La estrella azul B gira en una elipse kepleriana alrededor de la estrella A señalada en rojo, que ocupa un foco de la elipse. Obsérvese que la proyección de A no coincide con el foco de la órbita aparente.

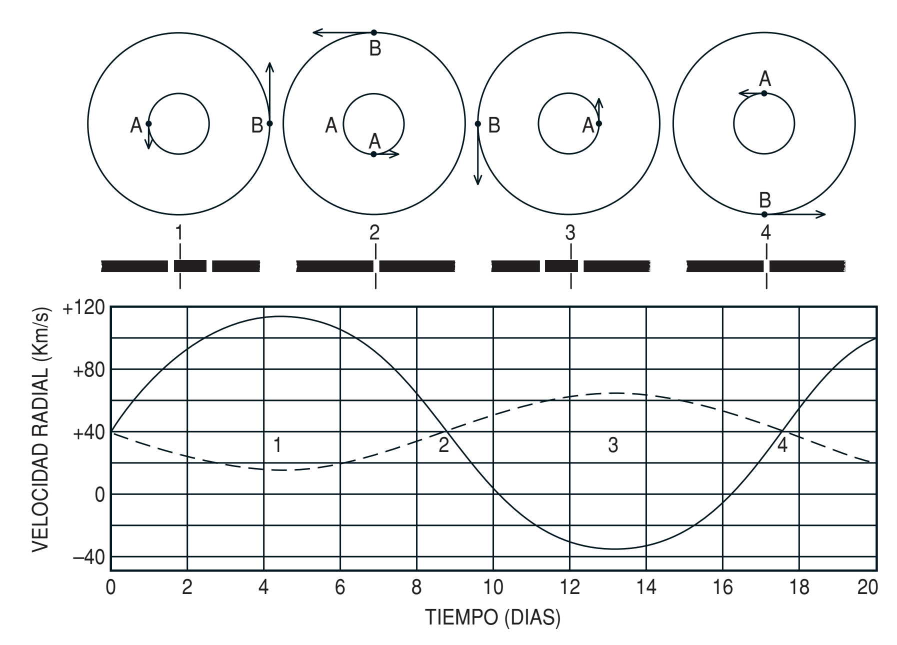
La investigación de los sistemas binarios de estrellas constituye la aplicación de un problema clásico de mecánica: el problema de los dos cuerpos. Los grandes astrónomos que acometieron su solución no encontraban ninguna razón para no tratar las estrellas binarias como dos puntos de luz, e identificaron sus velocidades orbitales con las velocidades de un movimiento kepleriano clásico.

En la mayoría de las binarias visuales, bien separadas, tal representación resulta adecuada. Se pueden considerar a sus componentes estrellas esféricas, con una distribución de la luminosidad sobre sus discos aparentes simétrica en torno a sus centros. En estas condiciones, podemos identificar las velocidades radiales medidas espectroscópicamente con las velocidades de dos esferas que describan un movimiento orbital que obedezca las leyes de Kepler; identificación que extenderemos, en el caso de las eclipsantes, a las variaciones fotométri-



cas medidas cuando un componente eclipsa al otro.

En realidad, las binarias espectroscópicas y eclipsantes están constituidas casi siempre por dos estrellas cercanas cuyas dimensiones no son despreciables en comparación con la distancia que las separa (son binarias "cerradas"). Aparecen entonces los "efectos" de proximidad. Estos efectos, debidos a las interacciones entre dos estrellas muy próximas, se detectan y estudian sobre todo en las binarias eclipsantes; han proporcionado el conocimiento fundamental de la estructura física de estrellas de diferentes tipos.



5. CURVAS DE VELOCIDAD RADIAL de una hipotética binaria espectroscópica con órbitas circulares. Al acercarse o alejarse las estrellas componentes en direcciones opuestas, el astrónomo observa cómo las líneas del espectro se desplazan por efecto Doppler a un lado u otro (1 y 3) respecto a la posición en que mantienen constante su distancia al observador (2 y 4).

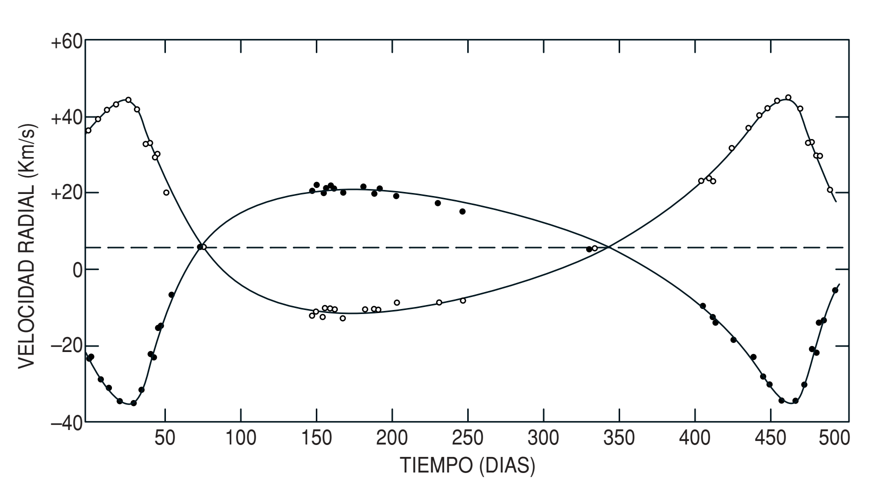
Binarias eclipsantes

En 1782 John Goodricke, con sólo 18 años, acertó a explicar correctamente las variaciones periódicas del brillo de la estrella β Persei o Algol, en la constelación de Perseo. Esta estrella de segunda magnitud, perceptible a simple vista, representó un enigma ya para los chinos y árabes medievales, que habían observado que su luz disminuía un tercio cada dos días y medio. Los árabes la llamaron Al Ghul, o estrella del demonio, de donde viene el nombre de Algol. Montanari había descrito en 1671 las curiosas oscilaciones de Algol; las observó desde Bolonia. Pero hasta Goodricke nadie pensó que pudiera tratarse de una binaria eclipsante, es decir, de una estrella que era eclipsada periódicamente por otra más oscura que giraba a su alrededor.

El descubrimiento de Goodricke de la primera binaria eclipsante no se aceptó fácilmente. Hasta casi un siglo más tarde, en 1880, año en que Edward C. Pickering estudió Algol, nadie tomó en serio la existencia de binarias eclipsantes. En 1912 Henry Norris Russell y Harlow Shapley desarrollaron el primer método matemático para analizar las *curvas de luz* (variaciones luminosas de las binarias eclipsantes en función del tiempo). Lo aplicaron a cualquier tipo de eclipse (total, parcial o anular). Hoy día se conocen cientos de binarias eclipsantes, aunque no todas se han estudiado con precisión.

El análisis de las curvas de luz de las eclipsantes permite determinar la forma de las órbitas y su tamaño, la relación de luminosidad de las dos estrellas componentes, variaciones de luminosidad en el disco de las estrellas (oscurecimiento en el limbo) y efectos de proximidad (por ejemplo, la deformación de la forma esférica debida a fuerzas de marea, el intercambio de materia entre las componentes y los efectos de reflexión de que se hablará más adelante). Proporcionan, además, los mejores datos que tenemos sobre atmósferas estelares extensas. Pero, sin duda, los datos más importantes derivados de las curvas de luz se refieren al tamaño de las estrellas en relación con su distancia mutua, su temperatura superficial y la inclinación *i* de las órbitas respecto a la visual, necesaria para determinar las masas estelares en las binarias espectroscópicas.

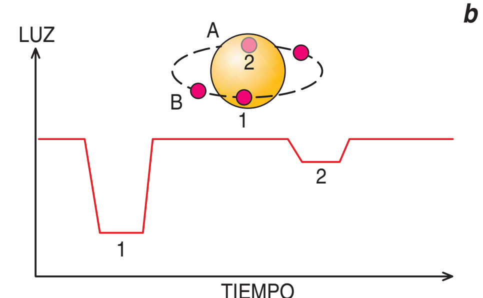
Las curvas de luz y las paralajes permiten deducir, por otra parte, las dimensiones de las órbitas y la

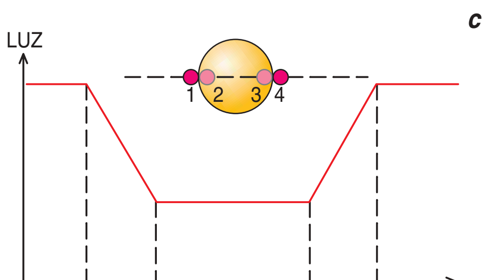


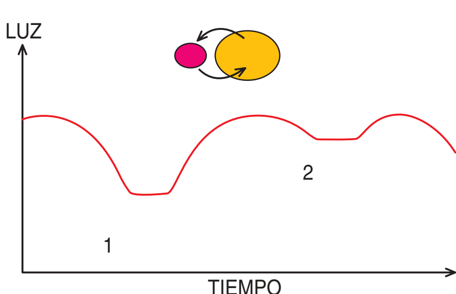
6. CURVAS DE LAS VELOCIDADES RADIALES de las dos estrellas que integran la binaria espectroscópica ϕ Cygni.

TEODORO VIVES (*arriba*); TEODORO VIVES, ADAPTADO DE RACH Y HERBIG (*abajo*)









7. CURVAS DE LUZ de binarias eclipsantes. En la parte 1 del esquema básico (*a*) se produce el eclipse principal: la estrella menos brillante B eclipsa a A; en el secundario, en 2, ocurre lo contrario. Entre eclipses, la curva representa la suma de la luz de las dos estrellas. En un eclipse central (*b*), la estrella menor (*roja*) se oculta detrás de la mayor (*amarilla*), recorriendo su diámetro. El tamaño de las estrellas se determina midiendo el tiempo entre los contactos (estrella menor entre 1 y 2 o 3 y 4; la mayor entre 1 y 3 o 2 y 4), si se conoce por el espectro la velocidad re-

velocidad con que las recorren las estrellas. Si los eclipses son centrales, es decir, si la inclinación *i* es cero y el plano orbital se ve exactamente de canto, basta medir el tiempo que duran los eclipses para obtener el tamaño de las estrellas.

En el eclipse central, una de las estrellas recorre con velocidad conocida el diámetro de la otra; la duración del eclipse permite deducir en ese caso la longitud del diámetro, es decir, el tamaño de la estrella. Cuando los eclipses no son centrales, el problema geométrico se torna algo más complicado, pero existen métodos para resolverlo. Se ha descubierto así que existen estrellas de tamaño muy diferente, desde estrellas enanas, como las estrellas de neutrones, de unos 10 kilómetros de diámetro, o las enanas blancas, del tamaño de la Tierra, hasta supergigantes, como las componentes principales de las binarias eclipsantes ζ y ε Aurigae, cuyos diámetros superan el del Sol unas 490 y 2556 veces, respectivamente.

La historia de la determinación de los parámetros físicos de las binarias eclipsantes, a partir de sus curvas de luz, se puede dividir en tres períodos: primero se idearon métodos geométricos, después métodos analíticos, basados por lo general en ajustes por mínimos cuadrados, y finalmente modelos físico-matemáticos computables.

El *problema directo* consiste en calcular la luz que un observador distante recibirá del sistema binario a partir de elementos dados: la forma, posición y distribución de la luminosidad aparente sobre el disco de cada estrella. El *problema inverso* estriba en encontrar los elementos del sistema a partir de los valores de la curva de luz. En el caso ideal, el número de valores requeridos coincide con el de elementos a determinar. Si tales valores han sido calculados como solución del problema directo, la solución del problema inverso reproducirá los elementos iniciales exactamente. Si los valores han sido obtenidos mediante observaciones astronómicas, los errores de las observaciones afectarán a los elementos y la solución no será nunca exacta. Surge entonces el *problema observacional*, que consiste en obtener, a partir de numerosas observaciones fotométricas afectadas de error, los elementos que mejor las representan.

Modelos

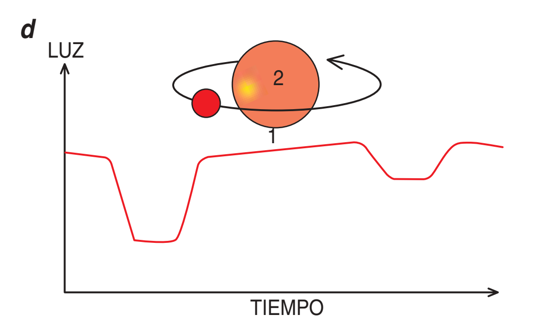
Russell y Shapley aportaron en 1912 el primer modelo geométrico de solución tanto del problema directo como del inverso: el modelo esférico. En él, dos estrellas esféricas se mueven en órbitas circulares según las leyes de la mecánica; sus discos están oscurecidos hacia el limbo de acuerdo con una ley del coseno del ángulo entre el radio de cada estrella y la línea de la visual. Esta ley es válida en primera aproximación para nuestro Sol, cuyo disco aparece más oscuro y rojizo hacia el borde o limbo. En el centro del disco solar podemos mirar hasta cierta profundidad, que corresponde a una capa fotosférica con determinada temperatura; en cambio, cerca del borde vemos una capa más profunda, más superficial y fría, que emite menos luz y más rojiza. En las binarias eclipsantes el oscurecimiento en el limbo influye en la forma de los mínimos de luz.

La solución del problema observacional con el modelo esférico se obtiene en primer lugar por métodos gráficos; sólo si la precisión de las observaciones lo justifica, la etapa final será una solución definitiva por mínimos cuadrados. No hay que perder nunca de vista, en efecto, que el grado de aproximación con que se pueden obtener los parámetros de una binaria eclipsante depende de la bondad de las observaciones. El método estadístico de los mínimos cuadrados sólo se puede aplicar a las mejores observaciones, ya que el trabajo laborioso que requiere no está justificado más que en esos casos. Por otra parte, no es un algoritmo mágico que proporcione siempre la solución correcta. La existencia de errores sistemáticos en las observaciones, cuya presencia se puede comprobar con harta frecuencia al comparar los

El autor

TEODORO VIVES SOTERAS ha desarrollado su actividad en diferentes institutos y observatorios de Europa y América. Profesor de astronomía y astrofísica, pertenece a la Unión Astronómica Internacional. Codirige el Observatorio de Calar Alto en Almería.

TEODORO VIVES



lativa de las dos estrellas. Las fuerzas gravitatorias deforman las componentes muy próximas de una binaria eclipsante cerrada; la distorsión afecta a la curva de luz (*c*). El efecto de reflexión —el calentamiento local de la estrella más fría por la más caliente— también la altera (*d*).

diferentes valores obtenidos por distintos observadores, falsea en su raíz el empleo de mínimos cuadrados.

Se trata, además, de ecuaciones trascendentes no lineales (es decir, sus soluciones no son funciones polinómicas); el método de mínimos cuadrados, pues, no podrá resolverlas directamente. Hay que lograr una solución preliminar por métodos gráficos o analíticos directos, y aplicar a continuación el método de mínimos cuadrados a las correcciones diferenciales de esa solución preliminar. Así han procedido durante muchos años numerosos investigadores de las curvas de luz de las binarias eclipsantes. Sin embargo, en las últimas décadas, el estudio de los sistemas físicos complejos no lineales, que ha dado lugar al nacimiento de la *teoría del caos* determinista, arroja serias dudas sobre esa estrategia de investigación. Hoy día sabemos que todos los sistemas mecánicos —entre ellos las estrellas binarias, a las que además perturban gravitatoriamente otros cuerpos— son sistemas caóticos que en rigor no admiten una representación cabal a través de simplificaciones lineales.

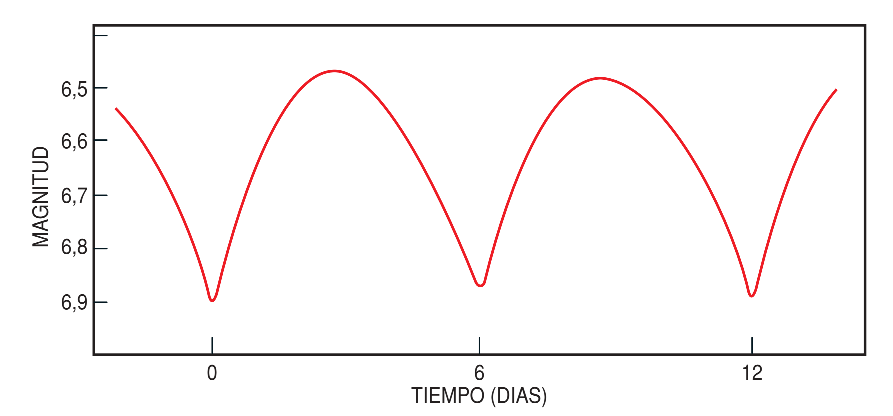
La imposibilidad de resolver directamente las ecuaciones básicas no lineales de una binaria eclipsante, tanto con métodos gráficos como analíticos, ha obligado a emplear procedimientos iterativos o por aproximaciones sucesivas. De ese carácter iterativo es el método gráfico inicial de Russell y Shapley, llevado a un alto grado de refinamiento con la publicación en 1953 de los nomogramas de J. E. Merrill. Posteriormente, Z. Kopal desarrolló un método numérico iterativo que no requiere otra interpolación gráfica que la necesaria para determinar la profundidad de los mínimos en las curvas de luz y los tiempos de contacto de los eclipses.

Cuando las dos componentes de un sistema eclipsante están separadas por menos de diez radios, las fuerzas gravitatorias y centrípetas deforman las estrellas. Se producen entonces sobre sus superficies variaciones de luminosidad, engendradas por las variaciones de la gravedad y el desigual calentamiento de cada una al recibir la radiación de la otra: es el efecto de reflexión. El sistema se representa mediante el *modelo elipsoidal*, que asimila las estrellas a elipsoides de tres ejes. La falta de esfericidad hace que el sistema aparezca más brillante en las cuadraturas —cuando las dos estrellas y la Tierra forman un triángulo rectángulo— que inmediatamente antes o después de los eclipses. El efecto de reflexión produce, en caso de que las dos estrellas difieran apreciablemente en luminosidad, más brillo cerca del eclipse secundario —el de la estrella menos brillante por la más brillante— que del primario.

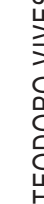
Rectificación

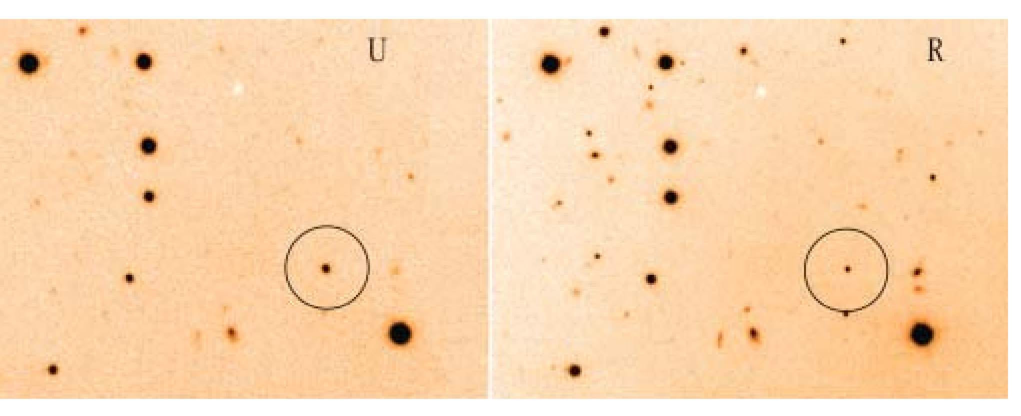
En teoría, la representación exacta de la curva de luz de un sistema elipsoidal no es posible con un modelo geométrico simple. Por eso se viene recurriendo tradicionalmente al artificio denominado *rectificación*. La rectificación consiste en eliminar de la curva de luz correspondiente al modelo elipsoidal los efectos de la no esfericidad y de reflexión, transformándola en una curva *rectificada* que corresponde al modelo esférico. Las constantes necesarias para la rectificación se pueden hallar gracias a la variación fuera de los eclipses. Una vez derivados los elementos de la curva rectificada se pueden determinar los del modelo rectificable o elipsoidal.

Todos estos métodos tradicionales de análisis de curvas de luz son tremendamente complicados y los modelos en que se basan adolecen de limitaciones fundamentales, algunas de las cuales se han apuntado ya. No es posible determinar, antes de obtener elementos al menos provisionales, cuál de dos mínimos consecutivos de una curva de luz corresponde al eclipse de la estrella menor por la mayor (ocultación) o al eclipse de la mayor por la menor (tránsito). Por otra parte, el análisis de los elementos geométricos es diferente en cada tipo de eclipse (total, parcial o anular) con la consiguiente complicación en el análisis. Además, no podemos estar nunca completamente seguros de que el efecto de una mayor o menor proximidad en cada momento haya sido eliminado de forma satisfactoria por la rectificación ni, por consiguiente,



8. CURVA DE LUZ en el visual de la binaria eclipsante V453 Sco, obtenida por el autor de este artículo a partir de más de mil observaciones fotoeléctricas en el Observatorio Interamericano de Cerro Tololo (Chile). Se trata de una eclipsante del tipo β Lyrae con un período de 12 días y una variación luminosa máxima de 0,4 magnitudes estelares. Los mínimos principales son tan sólo una centésima de magnitud más profundos que los secundarios.



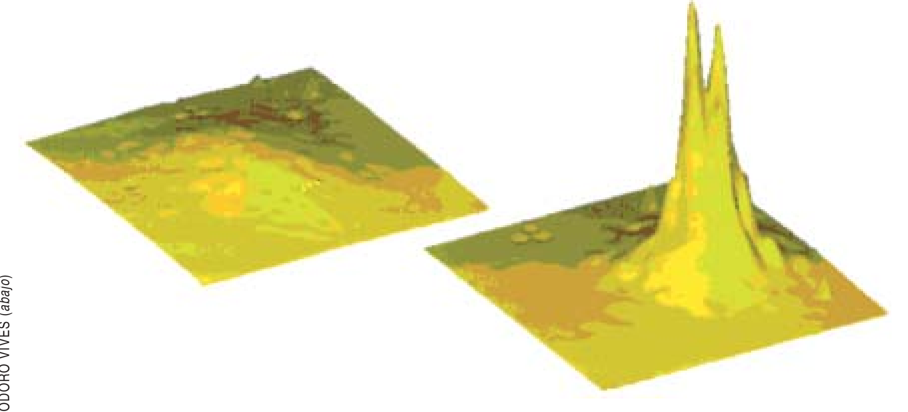


9. RX J0806.3+1527 es el sistema binario observado con el período de rotación más breve: 321 segundos. Sus dos componentes giran, una alrededor de la otra, a unos 1000 km/s. Se conocía a este objeto como una fuente de rayos X variable; en 2002 se descubrió que está constituido por dos enanas blancas separadas tan sólo 80.000 kilómetros. Estas fotografías se tomaron con filtros distintos.

de que la curva de luz rectificada represente únicamente las variaciones debidas a los eclipses.

Para obviar estas imperfecciones en el proceso de rectificación, Kopal propuso en 1961 no abordar las curvas de luz en función del tiempo, sino de la frecuencia: de las periodicidades que componen la curva de luz observada. La idea está tomada de la teoría de la información a la que recurren los ingenieros de telecomunicaciones. La variación de la curva de luz en función del tiempo de una eclipsante cerrada representa la superposición de los efectos fotométricos de proximidad, que son continuos, sobre el efecto de los eclipses, que son discontinuos. Además, la deformación de las dos estrellas produce un aumento de luminosidad de cada una de ellas por separado, mientras que el efecto de reflexión se suma a la luz total del sistema, sin que sea posible discriminar unos efectos de otros.

Estas dificultades desaparecen si se aplica la transformada de Fourier a la curva de luz, que transforma lo discontinuo en el tiempo en continuo en el dominio de las frecuencias y viceversa. Los mínimos en la curva de luz debidos a los eclipses, que son discontinuos en función del tiempo, se transforman en una función continua de la frecuencia y, lo que reviste mayor importancia, el espectro de frecuencia de todos los efectos de proximidad, continuos respecto al tiempo en la curva de luz, resultan discretos y su distribución se puede determinar previamente. Por consiguiente, la separación de los efectos de proximidad y de los eclipses resulta automática en el dominio de las transformadas de Fourier; la rectificación es innecesaria.



10. COMPARACION DE IMAGENES. A la izquierda, imagen obtenida directamente con el telescopio de 3,5 metros del Observatorio de Calar Alto de Almería en la longitud de onda de 2,2 μm, apenas una mancha difusa, de la estrella binaria φ UMa. A la derecha, la imagen de la misma estrella claramente resuelta en sus dos componentes, distantes 0",24, corregida con el sistema ALFA de óptica adaptativa de dicho telescopio.

GIAN LUCA ISRAEL *et al.*, OBSERVATORIO EUROPEO DEL SUR, PR Photo 10a/02 (*arriba*); TEODORO VIVES (*abajo*)

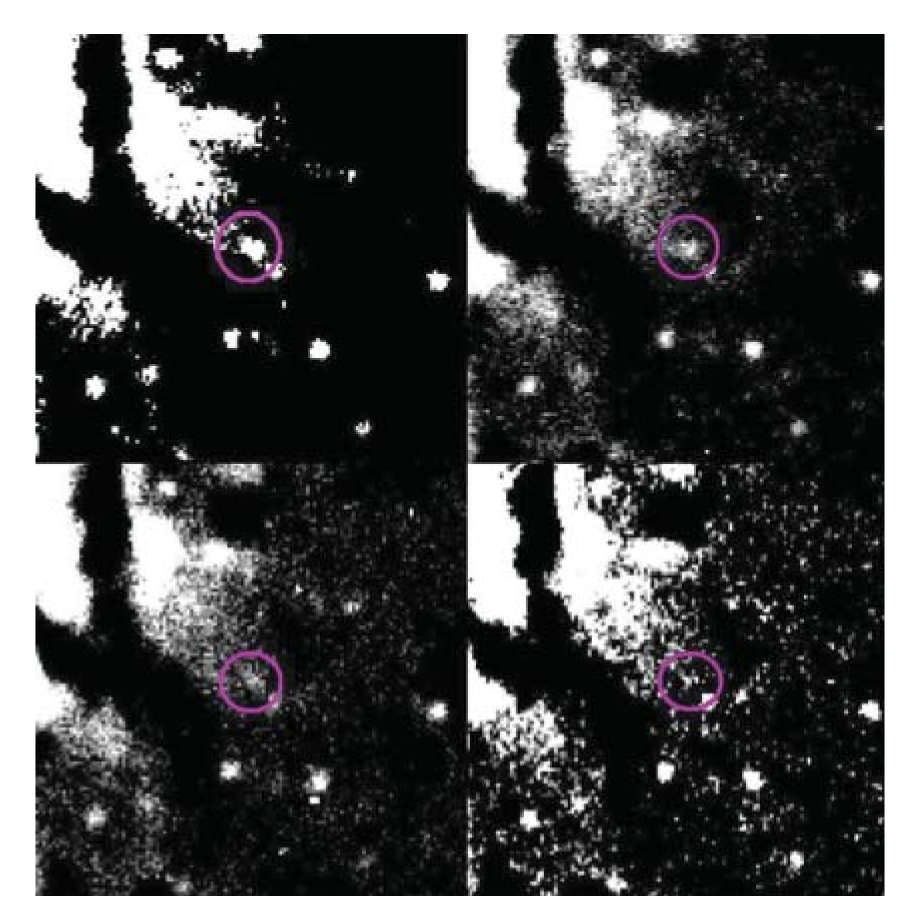
Simulación por ordenador

En los últimos años, el avance de los ordenadores rápidos ha permitido que numerosos investigadores construyan curvas de luz sintéticas a partir de modelos físicos que incluyen los parámetros característicos de las binarias eclipsantes. Los modelos más conocidos se deben a S. M. Rucinski, R. E. Wilson, D. B. Wood, G. V. Cochran o G. Hill y J. B. Hutchings. Introducido el programa del modelo, el ordenador va calculando una serie de curvas de luz con diferentes valores específicos de los parámetros que lo definen: radios estelares, masas, temperaturas efectivas y coeficientes de oscurecimiento de las estrellas, características de las órbitas e inclinación de las mismas, etc. La curva de luz observada de una determinada binaria eclipsante se compara con las curvas calculadas hasta conseguir el mejor ajuste. Los parámetros de la curva sintetizada por el ordenador que mejor se ajusta se adoptan entonces como solución del sistema binario.

Este juego de simulación informática de curvas de luz de estrellas binarias eclipsantes puede resultar divertido, a pesar de su complejidad, pero lo que no está tan claro es que sea realmente fiable. Porque de hecho se puede obtener la misma curva, la más ajustada, con diferentes conjuntos de parámetros. Basta aumentar los valores numéricos de unos y disminuir los de otros. En otras palabras, no hay una solución única: es posible reproducir con el ordenador la curva observada introduciendo estrellas con diferentes parámetros relativos a temperaturas, tamaños o inclinaciones orbitales. Pero la unicidad de la solución no agota todos los problemas inherentes a los modelos computables. La construcción del modelo mismo es objeto de debate

entre los investigadores; existe una viva controversia entre los especialistas sobre si es mejor un modelo tan complicado como sea posible, que incluya todos los efectos imaginables, o construir un modelo simple y tratar los efectos como perturbaciones.

El investigador de estos difíciles problemas de astrofísica, en los que interviene necesariamente un complejo modelo físico-matemático, no debe dejarse deslumbrar por algoritmos y desarrollos matemáticos formidables. Tampoco debe dejarse arrastrar por la magia de los ordenadores que, en fracciones de segundo, resuelven complicadas ecuaciones e imprimen vertiginosamente soluciones, con el único defecto de tener tal vez poco que ver con la realidad del mundo físico. Los ordenadores son poderosos e imprescindibles instrumentos en la investigación científica moderna, pero sólo cuando siguen y obedecen a la mente crítica del investigador, cuya única garantía de objetividad es el contraste de los modelos con las observaciones y el rigor lógico, y no al revés.



11. LA EXPLOSION DE UNA NOVA en la galaxia NGC 1316, del cúmulo de Fornax, tal y como fue vista los días 9, 12, 15 y 19 de enero de 2000 (*de arriba abajo y de izquierda a derecha*); la diferencia de brillo salta a la vista. Este tipo de explosiones se producen de ordinario en sistemas binarios, donde uno de los componentes es una enana blanca. A partir de su brillo aparente se calibran distancias. Esta se halla a 70.000.000 de años-luz: se trataba, junto con otras tres observadas en la misma galaxia y en el mismo período por el Observatorio de Paranal, en Chile, de la más lejana encontrada hasta la fecha.

M. DELLA VALLE *et al.*, OBSERVATORIO EUROPEO DEL SUR, PR Photo 18c/00

Binarias y supernovas

El extraordinario desarrollo actual de nuevos instrumentos y técnicas de observación abre anchas perspectivas a la investigación de las binarias. El satélite alemán ROSAT de rayos X, por ejemplo, puesto en órbita en 1990, ha permitido recientemente localizar un curioso tipo de estrellas binarias de rayos X formadas por una enana blanca y una gigante roja.

La asociación de una enana blanca con una estrella común para formar una binaria se conoce desde hace tiempo. La más famosa de todas es Sirio, la estrella más brillante del firmamento nocturno, cuya compañera Sirio B fue vista ya en 1862 por Alvan G. Clark. En 1914 Walter S. Adams analizó el espectro de Sirio B y demostró que se trataba de una estrella luminosa, caliente, sorprendentemente pequeña y muy densa. Fue la primera enana blanca estudiada por los astrónomos. Hoy sabemos que abundan las enanas blancas; con cierto detenimiento se han estudiado unas 600. Pero las más interesantes para los astrofísicos son las asociadas estrechamente a estrellas gigantes (constituyen binarias de rayos X) o a estrellas de neutrones (dan origen a púlsares).

Las binarias de rayos X detectadas en los últimos años emiten radiación “dura” de alta energía, pero el satélite ROSAT ha observado ahora binarias de rayos X “blandos”, de menor energía. La radiación X, que no atraviesa la atmósfera y sólo se puede observar desde satélites, se produce al incidir materia de la estrella compañera sobre la enana blanca. Su acumulación sobre la enana blanca ha agotado ya en su interior el hidrógeno, que se fusiona en helio; desencadena en la superficie de ésta las reacciones nucleares que originan los rayos X blandos observados por el ROSAT. Según E. P. J. Van den Heuvel, Peter Kahabka y Saul A. Rappaport, el enorme interés de ese proceso reside en que explica la explosión de las *supernovas del tipo Ia* —las ligadas a enanas blancas cuyos espectros presentan silicio pero no hidrógeno—, utilizadas en cosmología para medir las distancias de las galaxias más lejanas. En definitiva, las binarias en que una estrella común orbita alrededor de una enana blanca constituyen un verdadero laboratorio de astrofísica donde investigar las propiedades de las estrellas; además, permiten afinar los modelos de las supernovas Ia.

También ofrecen espectaculares posibilidades a la investigación de las binarias los nuevos instrumen-

tos y métodos de los observatorios terrestres. Mencionaremos sólo la óptica adaptativa, que elimina de las imágenes telescópicas el influjo perturbador de la turbulencia atmosférica al rectificar las deformaciones de la onda luminosa con la ayuda de espejos "blandos" o deformables. La forma de éstos se modula a fin de compensar las pequeñas ondulaciones turbulentas sufridas por la luz.

En 1990, el telescopio de 3,6 metros de ESO en La Silla, fue, de entre los destinados a usos civiles, el primero equipado con óptica adaptativa. Le siguió en 1996 otro telescopio, también de 3,6 metros, del Observatorio de Mauna Kea, Hawai. Pero en los dos casos el sistema no disponía de estrella artificial generada por láser; había que limitar las observaciones a las proximidades de estrellas que brillasen tanto como para que se pudiese medir la turbulencia de la atmósfera. En el Observatorio de Calar Alto, del Centro Astronómico Hispano-Alemán de Almería, se instaló, en octubre de 1996, el sistema de óptica adaptativa ALFA, acrónimo de *Optica adaptativa con Láser para la Astronomía*, que dispone de una estrella artificial correctora, producida por un rayo láser en la alta atmósfera, a una altura de entre 90 y 100 km; es posible efectuar con él observaciones casi en cualquier dirección del cielo. La compensación de la turbulencia del aire por medio de la óptica adaptativa permite observar, con gran resolución, estrellas binarias débiles. Con sistemas así se augura un futuro prometedor para la observación de las estrellas binarias también desde los observatorios terrestres.

Bibliografía complementaria

THE BINARY STARS. R. G. Aitken, Dover Publications. Nueva York, 1964.

INTERACTING BINARY STARS. J. Sahade, F. B. Wood, Pergamon Press. Londres, 1978.

ASTRONOMÍA X. A. Giménez Cañete, A. Castro Tirado. Equipo Sirius. Madrid, 1998.

Artur Hahn

Física del giroscopio. ¿Cómo funciona el Gyrotwister?

¿Una paradoja?

"Powerball", "Gyrotwister" o "Rolerball" son los nombres con los que se vende un artilugio que entrena los músculos de la mano. Se trata de un giroscopio esférico. Rota suspendido en el interior de una ligera, delgada y traslúcida esfera hueca; a través de la misma se observa el cuerpo que lleva dentro. El giroscopio queda suspendido gracias a un eje de pequeño diámetro, cuyos extremos se insertan en una ranura de la cara interna de la cápsula. La ranura abarca un círculo máximo entero de la cápsula; es un poco más ancha que el diámetro del eje, para que éste pueda girar libremente o rodar por los laterales de la ranura. Se puede acceder al giroscopio a través de un agujero circular abierto en la cápsula, de manera que se le podría hacer girar con el pulgar, o tirando de una cuerda unida a su ecuador.

El juego consiste en acelerar el giroscopio, sin tocarlo directamente, hasta un número de revoluciones mucho mayor. Hay que mover el artilugio con habilidad en la mano.

¿Cómo llega a acelerarse? El giroscopio va unido a la cápsula sólo por los dos puntos en los que su eje toca la cara interna de la ranura. Las fuerzas que se puedan transmitir por esos puntos llevan irremediablemente la dirección equivocada, pues el momento angular que generan será perpendicular al eje del giroscopio: a lo sumo cambiará la dirección en que gire éste, pero la rotación no se acelerará.

O, al menos, éste era el resultado al que llegamos, una noche, entre amigos, al aplicar los principios de la física clásica. Habíamos estado buscando la aceleración que da sentido al aparato. Al final lo conseguimos. Refutamos la teoría con la práctica. ¿Dónde estaba nuestro error de razonamiento?

Formas de moverse del giroscopio

Si se tiene en la mano el aparato mientras gira deprisa, se verá que el giroscopio rota alrededor de un eje de rotación que no cambia con el tiempo; el número de rotaciones por segundo irá poco a poco disminuyendo. La interacción con la cápsula se restringe a cierta fricción entre el eje del giroscopio y una posición (siempre la misma) de la ranura-guía. Este es el movimiento ordinario de un giroscopio. No cuesta desplazar el eje de rotación paralelamente, pero cambiar su orientación en el espacio requiere un considerable esfuerzo, tanto mayor cuanto más deprisa gire el giroscopio.

La rotación se caracteriza mediante una magnitud vectorial, el momento angular. Para un giroscopio esférico, ese vector reposa en el eje instantáneo de rotación y apunta en el sentido en que se enroscaría un tornillo que girara como el giroscopio. El módulo del momento angular es igual, en determinados casos especiales, el giroscopio esférico entre ellos, al producto de la velocidad angular y del momento de inercia respecto al eje de giro; el momento de inercia crece con la masa que gire y su distancia al eje de rotación.

Para el momento angular valen principios análogos a los del momento a secas, o producto de la masa por la velocidad, que se conserva siempre que no actúe ninguna fuerza. El momento angular de un cuerpo en rotación cambia sólo si actúa sobre él un momento de fuerza *M;* la variación (derivada temporal) del momento angular es igual al momento de fuerza que intervenga: $d\vec{L}/dt = \vec{M}$ o $d\vec{L} = \vec{M}\,dt$. Hay una definición breve de *M*: "brazo de palanca por fuerza" ($\vec{M} = \vec{R} \times \vec{F}$). Ahí "por" se refiere al producto vectorial de dos vectores: el vector $\vec{M}$ es perpendicular a los vectores $\vec{R}$ y $\vec{F}$, y su módulo (su "longitud") coincide con el producto ordinario de los módulos de $\vec{R}$ y $\vec{F}$ sólo si estos vectores son perpendiculares entre sí; si no, será menor.

Sobre nuestro tranquilo giroscopio no actúan, casi, momentos de fuerza. Su momento angular permanece



1. El Gyrotwister se recomienda a deportistas, pianistas y usuarios de teclados de todo tipo para fortalecer la musculatura de la mano

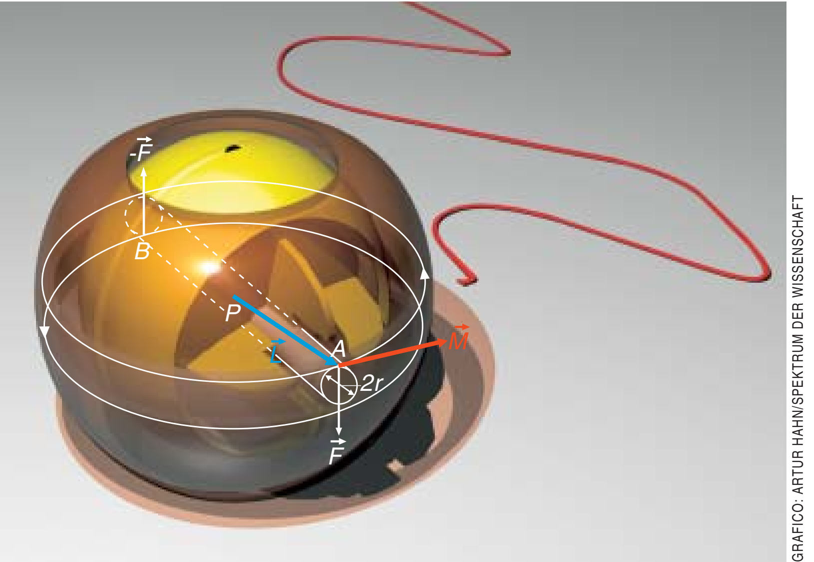


GRAFICO: ARTUR HAHN/SPEKTRUM DER WISSENSCHAFT

2. Mientras los márgenes de la ranura-guía obligan al eje del giroscopio "a seguir la pista", ejercen un par de fuerzas ($\vec{F}$, $-\vec{F}$) sobre los puntos A y B. El momento de fuerza que se genera es perpendicular a $\vec{F}$ y al vector $\vec{R}$ del brazo de palanca, que va de P a A. Como el eje instantáneo de rotación, y por ello el vector del momento angular, apuntan en la misma dirección que $\vec{R}$, el momento angular y el momento de fuerza serán perpendiculares entre sí

constante (aproximadamente) en dirección y módulo. Y aunque actuara un momento de fuerzas cuyo vector fuera perpendicular al vector del momento angular, sólo cambiaría la dirección de éste, pero no su módulo (*dibujo de la derecha, abajo*).

Precesión sin empuje

Más interesante que esta rotación libre es el movimiento de precesión. Sujetemos con fuerza la cápsula de tal manera que no se mueva (con las grandes fuerzas que aparecen, lo conseguiremos sólo en cierta medida). La ranura que sirve de guía en la cápsula quedará quieta, pues, en el espacio. El eje del giroscopio se irá desplazando ("precesionará") lentamente por el círculo que le marca la ranura. Al contrario de lo que sucede en la rotación libre, no rozará la cara interna de la ranura, sino que rodará sin deslizamiento por ella (*véase la figura 2,* donde los márgenes de la ranura que figuran en el dibujo pueden hallarse en cualquier posición), a un lado por el margen "superior", al otro lado por el "inferior".

El diámetro de la cápsula y, por tanto del margen de la ranura, es más o menos cincuenta veces mayor que el diámetro del eje del giroscopio allí donde toca el margen: si observamos con mayor detenimiento, veremos que el eje se reduce en sus extremos hasta un grosor de un milímetro, mientras que el juguete entero tiene más o menos unos cinco centímetros de diámetro. Por tanto, cuando el giroscopio complete cincuenta vueltas, habrá rodado una vez por el perímetro (interno) de la cápsula: la frecuencia de precesión es a la de rotación como el radio del eje giroscópico al de la cápsula; en este caso, 1 a 50.

No intervienen más fuerzas que las que ejercen los márgenes de la ranura sobre los dos extremos del eje del giroscopio. El momento de estas fuerzas es perpendicular al eje del giroscopio. Se debe distinguir entre el eje (material) del giroscopio y el eje instantáneo (infinitamente delgado) de rotación, esto es, la línea recta cuyos puntos tienen en cada instante velocidad cero. Coincide con la recta que pasa por los puntos de contacto *A* y *B*; en general cambia con el tiempo. Pero el vector del momento de fuerzas generado por los márgenes de la ranura está dispuesto perpendicularmente al eje instantáneo y, por tanto, también al momento angular: de éste sólo se alterará la dirección (lo que sucede, pues el giroscopio precesiona), pero no su módulo. En particular, el número de revoluciones, que es proporcional al módulo del momento angular, no puede aumentar.

Pasado un corto tiempo *dt*, el punto de contacto *A* habrá recorrido un fragmento del margen o círculo superior (*véase la figura 2*); el punto de contacto *B* se habrá desplazado el mismo trecho en el círculo inferior. Como resultado, el eje instantáneo de rotación, y con él el vector del momento angular, habrá trazado un cono —muy plano—, con el vértice en el baricentro *S*.

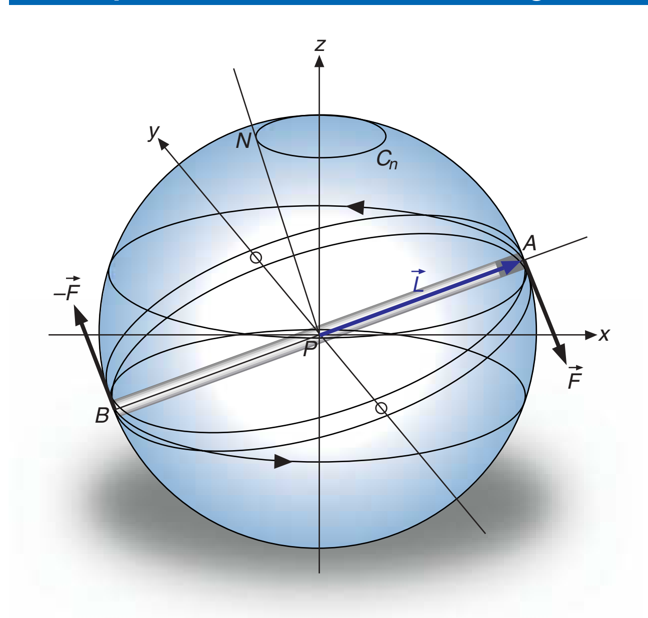
En la práctica, no se podrá inmovilizar el juguete en la mano. Si ésta se limita a reaccionar con su flexibilidad a las fuerzas de reacción del giroscopio, sin darle ningún impulso intencionadamente, actuará como un sistema de muelles del que estuviese colgado el juguete. Giroscopio y muelles pueden intercambiar energía; lo mismo que cuando unas fuerzas giroscópicas tensan un muelle, que al destensarse devuelve al giroscopio la energía que se le transmitió. Pero así no puede aumentar la energía total del sistema, y con ella el número de revoluciones del giroscopio. Durante ese movimiento el momento de fuerzas permanece perpendicular al momento angular (*véase el recuadro* " El caprichoso movimiento del giroscopio").

3. Si la variación $d\vec{L}$ del momento angular $\vec{L}$ es perpendicular al propio $\vec{L}$, de éste sólo variará la dirección (izquierda). *Su módulo sólo crecerá cuando $d\vec{L}$ tenga una componente en dirección de $\vec{L}$* (derecha)

Empuje

Si se desea que el giroscopio dé más vueltas por segundo, se le tendrá que proporcionar energía, es decir, realizar un trabajo. El trabajo es "fuerza por recorrido"; o más exactamente: fuerza por la componente del recorrido en la dirección de la fuerza. Pero sólo podremos ejercer una fuerza a través del punto de contacto del eje del giroscopio y el margen de la ranura, que hasta ahora era siempre un punto del eje ins-

El caprichoso movimiento del giroscopio



Por la cara interna de la cápsula, aproximadamente esférica, del giroscopio, a lo largo del ecuador, discurre la ranura-guía, representada en estas ilustraciones por un par de círculos poco separados. La perpendicular al plano de la ranura que pasa por el centro de la esfera atraviesa ésta en el "polo norte" *N*. Durante el movimiento de precesión libre de empuje, la cápsula y la ranura siguen el movimiento de precesión del giroscopio de suerte que el punto *N* trazará un pequeño círculo C_n alrededor del eje de precesión, fijo en el espacio, en la dirección de la flecha y con la frecuencia de precesión del giroscopio.

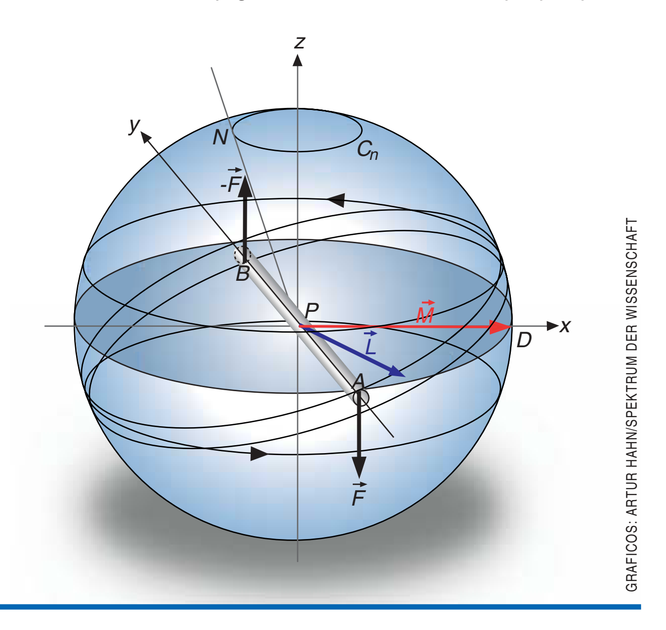
Añadamos un sistema de coordenadas rectangular de tal manera que el eje *z* coincida con el eje de precesión y que la línea de unión de los puntos de rotación *A* y *B* —y con ella, el vector del momento angular $\vec{L}$, así como los vectores de las fuerzas $\vec{F}$ y $-\vec{F}$— caiga en el plano *x-z*. Durante el movimiento de precesión, el dibujo entero girará en torno al eje *z*. En lo demás permanecerá inalterado. En particular, no cambiará la posición relativa de los vectores $\vec{L}$, $\vec{F}$, $\vec{M}$ y del "eje de la Tierra" *PN*. El punto *A* (como el *B*) no es un punto concreto del margen de la ranura; en un giro completo todos los puntos de ésta habrán desempeñado el papel del punto *A*. Pero el punto que aquí hace de punto *A* se halla en reposo, en ese instante representado, porque se encuentra en el punto superior de inversión del movimiento.

El momento de fuerza se orienta entonces en la dirección del eje *y* perpendicularmente al plano *x-z*. Es, en particular, perpendicular a la dirección del momento angular; el módulo de éste, pues, no variará.

Si, en cambio, para la misma posición de la cápsula, el eje de giroscopio no está en el plano *x-z*, sino en el eje *y* (*dibujo inferior*), el movimiento de precesión de la cápsula (que ya no se generará por sí solo, sino forzado por la mano del jugador) impartirá al giroscopio un momento de fuerza *M* gracias al par $\vec{F}$, $-\vec{F}$, un momento que no resultará perpendicular al vector del momento angular. Si *M* apunta en la dirección de *x*, en cambio, el vector del momento angular $\vec{L}$ no tendrá la dirección de la recta *PA*, sino aquella con la que el punto *A* —por efecto de la rotación alrededor del eje instantáneo de rotación, o sea, de la dirección de $\vec{L}$— tendrá la velocidad hacia abajo debida. ¡Así que $\vec{L}$ posee ahora una componente que no desaparece en dirección de $\vec{M}$! Por claridad, se ha exagerado la desviación del vector del momento angular $\vec{L}$ respecto al eje material del giroscopio.

¿En que consistía, pues, el error de razonamiento que cometimos al creer que no podía acelerarse el giroscopio? En confundir el eje material del giroscopio con el eje instantáneo de rotación.

Cuantitativamente: El giroscopio de un Gyrotwister tiene una masa de unos 250 gramos. De ello se obtiene (para una distribución de masa homogénea) un momento de inercia de 0,625 kg cm^2. El momento de fuerza que ejerce el juguete sobre la mano firme es proporcional al cuadrado del número de revoluciones. Para 200 vueltas por segundo (según las instrucciones se pueden alcanzar hasta 230) al menos 2 kg m^2/s^2. Así pues, el aparato, con un brazo de palanca de 2,5 cm, ejerce una fuerza de 80 N sobre la mano del jugador: ¡treinta veces su propio peso!



GRAFICOS: ARTUR HAHN/SPEKTRUM DER WISSENSCHAFT

tantáneo de rotación, un punto que, por definición, en el instante correspondiente, no se mueve. Donde no hay recorrido, no existe fuerza que ayude. Hemos de eliminar, pues, esta identidad entre punto de contacto y punto del eje de rotación.

Por consiguiente, debemos impartir al punto de contacto una velocidad distinta de cero. Para lograrlo de la manera más efectiva, se aplicará a la cápsula el mismo movimiento de precesión que se produciría en la precesión libre de empuje, pero con una fase adelantada 90 grados. Con el subsiguiente movimiento de la cápsula, el punto de contacto *A* se moverá hacia abajo, es decir, en la dirección de la fuerza eficaz: puesto que el empuje sobre el giroscopio será distinto de cero, su energía cinética, y con ella su velocidad de rotación, aumentarán.

Desentrañados los entresijos del juguete, el juego ha terminado.

JUEGOS MATEMÁTICOS

Juan M.R. Parrondo

La paradoja de la Biblioteca de Babel

El gran emperador de Babel ordenó a sus escribanos que caligrafiaran uno a uno los volúmenes de una grandiosa Biblioteca, una Biblioteca que contendría todos los libros de 500 páginas escritos y por escribir. Las instrucciones que dio al jefe de los escribanos eran muy sencillas. Deberían tomar las 27 letras del alfabeto, el espacio y los signos de puntuación "," "." ";" ":" "¡" "!" "¿" "?", y hacer todas las combinaciones posibles que se ajustaran a las 500 páginas de cada libro. Empezarían con un libro en blanco; después con un libro que sólo tuviera la letra "a" en su comienzo estando en blanco el resto. Otro con la letra "b", y así sucesivamente. Luego se empezaría con pares "aa", "ab", "ac", etc. El último de los volúmenes sería un libro cuyas 500 páginas consistirían en una sucesión interrumpida de signos de interrogación "?". Entre el primer libro en blanco y este último y misterioso libro lleno de interrogaciones estarían todos los libros escritos y por escribir. La mayoría serían ininteligibles o absurdos, pero en alguno de los estantes de la Biblioteca de Babel estarían *Don Quijote*, *La Crítica de la Razón Pura* o la exposición (sin fórmulas matemáticas) de la teoría de la relatividad.

El emperador estaba muy contento con su Biblioteca hasta que un joven estudiante pidió audiencia y le habló así:

"Majestad, habéis creado una gran obra con vuestra Biblioteca, puesto que encierra todas las verdades (y todas las mentiras), todas las historias y todas las bellezas del mundo que pueden expresarse con palabras. Pero hay algo en ello que no acabo de entender. Las instrucciones que disteis a vuestros escribanos para crear la Biblioteca se pueden dictar en menos de cinco minutos y cabrían en una simple hoja de papel. Esa hoja de papel contendría las instrucciones necesarias para crearla de nuevo. Si con unas instrucciones tan breves podemos crear todas las verdades (y todas las mentiras) del mundo, entonces cada una de estas verdades, o cada uno de esos libros, puede crearse con esas mismas instrucciones, o incluso con menos, puesto que cada libro es sólo una ínfima parte de la Biblioteca. Con su Biblioteca, Su Majestad ha demostrado que todas y cada una de las verdades y las historias de los hombres caben en una simple hoja de papel."

El emperador oyó al estudiante con atención, sin saber cómo responder. Tras despedirlo, se encerró durante semanas en los aposentos más retirados del palacio, sumido en una profunda confusión.

El argumento del joven estudiante parece absurdo, pero no puede descartarse fácilmente, y el Emperador lo sabía. Es cierto que las instrucciones para la creación de la Biblioteca caben en una hoja de papel. Por otro lado, como *El Quijote*, por ejemplo, es uno de los volúmenes de la Biblioteca, ¡está claro que sólo hace falta seguir esas instrucciones para escribirlo! ¿Cómo es posible?

Esta misma paradoja ocurre en un campo de las matemáticas, la *complejidad algorítmica*, creada por los matemáticos Kolmogorov, Chaitin y Solomonoff. Ya hablamos de ella cuando analizamos la relación entre la compresión de ficheros de datos y la teoría de la información (*Juegos matemáticos* de septiembre de 2001). La complejidad algorítmica de una cadena de símbolos es la cantidad de instrucciones que hay que dar a una persona o a un ordenador para que reconstruya dicha cadena. Si la cadena es muy grande pero responde a una determinada pauta lógica, su complejidad algorítmica es pequeña. Por ejemplo, la complejidad algorítmica del número de dos mil dígitos:

$$27272727\ldots27$$

es muy pequeña. Porque bastan las instrucciones "escriba 27 mil veces seguidas" para reconstruirla. Sin embargo, el número:

$$1222112121221112122\ldots$$

que se construye tirando una moneda dos mil veces y poniendo un 2 si sale cara o un 1 si sale cruz, tiene una complejidad algorítmica muy grande, ya que si los resultados del lanzamiento de la moneda son realmente aleatorios, lo más probable es que no tengan pauta alguna y las únicas instrucciones que permitan reconstruir ese número en concreto consistan en indicar al sujeto que quiera reconstruirlo todos los dígitos, uno por uno, que lo forman. En otras palabras, el número es incompresible: la información que contiene no puede reducirse a unas instrucciones más breves que el propio número; la única forma de describirlo es detallar, una por una, todas las cifras que lo forman.

La paradoja a la que se enfrenta el emperador de Babilonia, expuesta en términos de la complejidad algorítmica de Kolmogorov, Chaitin y Solomonoff sería la siguiente:

¿Cómo es posible que la complejidad algorítmica de una secuencia sea muy pequeña y que, sin embargo, la complejidad de un subconjunto de dicha secuencia sea grande? O: ¿cómo es posible que la complejidad algorítmica de la Biblioteca sea pequeña mientras que la complejidad de uno de sus volúmenes, *El Quijote* por ejemplo, es grande?

Para ver cómo y por qué es esto posible, consideremos la cadena C de dígitos formada por la concatena-

tenación de todos los números entre 0 y $10^{2000} - 1$. Para construir esta cadena *C* escribimos primero dos mil ceros, luego 1999 ceros y un 1, luego 1999 ceros y un 2, y así sucesivamente. Es decir, concatenamos todos los números de dos mil cifras, empezando por el cero, el 1, el 2, etc., hasta llegar al $10^{2000} - 1$. Esta inmensa cadena contiene

$$2000 \times 10^{2000} = 2 \times 10^{2003}$$

dígitos, pero su complejidad algorítmica es muy pequeña. Ahora bien, en la cadena *C* aparecen como subcadenas todos los números entre cero y $10^{2000} - 1$, y alguno de estos números —como el aleatorio que hemos visto antes— puede tener una complejidad algorítmica mucho mayor. Nos encontramos de nuevo con la paradoja de la Biblioteca de Babel.

La solución de la paradoja es la siguiente. Piensen en el número aleatorio de dos mil cifras que hemos dado antes. Este número —llamémoslo *n*— está dentro de la cadena *C*. Podríamos utilizar la pequeña complejidad algorítmica de *C* para describir *n*. Las instrucciones para reconstruir *n* serían: (*a*) constrúyase *C* y (*b*) encuéntrese *n* dentro de *C*. Sabemos que las instrucciones para llevar a cabo el paso (*a*) son muy breves. Pero, ¿qué ocurre con el paso *(b)*? Dentro de la cadena *C*, el número *n* empieza en el sitio:

$$2000 \times n$$

Luego, para precisar con absoluta exactitud dónde empieza *n*, ¡tendríamos que describir con absoluta exactitud el propio número *n*! Con lo cual utilizar la cadena *C* es completamente inútil: no nos ahorra ninguna instrucción.

El emperador de Babilonia encontró esta solución a la paradoja, pero no por sí solo, reflexionando en su palacio. La encontró gracias a uno de sus bibliotecarios. Después de muchos días de aislamiento y de tristes pensamientos acerca de las palabras del estudiante, decidió consultar por primera vez algún volumen de la Biblioteca (no lo había hecho nunca hasta entonces). Se acercó a la gran puerta dorada y preguntó al bibliotecario:

—He oído hablar de un libro divertido y a la vez revelador. Lo escribirá un tal Cervantes dentro de unos siglos, y trata de las aventuras de un caballero andante.

El bibliotecario le respondió sonriente:

—Sí, Majestad. Conozco ese libro. El escribano que lo transcribió me habló muy bien de él y recuerdo perfectamente en qué pasillo de la Biblioteca se encuentra. Os indicaré cómo llegar hasta él. Tomad este pasillo central hasta que veáis un gran cartel con la letra "e". Allí deberéis girar a la derecha y caminar un trecho no muy largo, hasta llegar al pasillo presidido por un cartel con la letra "n". Caminad sólo unos metros, porque enseguida deberéis tomar a la izquierda el pasillo correspondiente al espacio en blanco. Luego el pasillo de la letra "u", después el de la "n", a continuación el del espacio de nuevo. Más tarde el de la "l", el pasillo de la "u", luego el de la "g", el de la "a" y el de la "r"...



Un sencillo programa de ordenador podría generar una biblioteca que incluyese todos los libros posibles. Pero en el listado de ese programa no podríamos leer El Quijote*. Ilustración tomada del CD-ROM* La Biblioteca Total, Viaje por el Universo *de Jorge Luis Borges, ©1996, Nicolás Helft & Weber-Ferro S.R.L.*

El bibliotecario se pasó semanas indicando al Emperador el sitio exacto donde se encontraba el libro. Cuando terminó su explicación, el Emperador sonrió. No necesitaba recorrer la tupida red de corredores que formaban la Biblioteca. Sentado delante del mostrador del bibliotecario, había conocido con todo detalle las andanzas de Don Quijote y Sancho y, al mismo tiempo, había encontrado la solución a la paradoja.

Es cierto que los libros de la Biblioteca se podrían ordenar de modo tal que la tarea de describir el lugar en donde se encuentra *El Quijote* no equivaliera a recitarlo letra por letra. En *El Quijote*, como en cualquier otro libro escrito en una lengua real, hay un cierto grado de redundancia. Por ejemplo, detrás de una "q" siempre aparecerá una "u", con lo cual el bibliotecario se podría ahorrar parte de las indicaciones. Redundancia significa que en la cadena de letras que forman *El Quijote* cada letra nueva no nos "pilla completamente de sorpresa" sino que es parcial o totalmente predecible, como una "u" detrás de una "q", o como la "i" y la "a" detrás de "redundanc". Esta predecibilidad o redundancia hace que el lenguaje no sea óptimo, en el sentido de expresar la mayor cantidad de información por letra. De hecho, los compresores de ficheros que utilizamos en nuestros ordenadores detectan esta redundancia para reducir el tamaño del fichero que ocupa un texto hasta en un 70 %, como ya vimos la sección de *Juegos Matemáticos* de septiembre de 2001. Sin embargo, sería probablemente imposible aprender un lenguaje que estuviera completamente exento de redundancia.

En cualquier caso, el emperador puede dormir tranquilo: la escasa complejidad algorítmica de la Biblioteca no disminuye un ápice el valor de todas las joyas que se esconden en ella, puesto que la aventura de encontrarlas es idéntica a la aventura de escribirlas.

Mark Fischetti

VISION NOCTURNA

Ver en verde

Los seres humanos no vemos en la oscuridad. Sin embargo, se cuentan por millones los que observaron en sus televisores, durante la guerra de Irak, cómo se movían en la negrura del desierto las figuras fantasmales de tanques, soldados y periodistas. Fue posible gracias a los tubos de visión nocturna integrados en las videocámaras, la misma técnica con que los amigos del aire libre localizan animales nocturnos, o boyas en las aguas en plena noche, y la policía vigila delincuentes agazapados a oscuras.

Por cerrada que sea la noche, las estrellas, la luna y las iluminaciones artificiales proporcionarán un reducido número de fotones. El tubo de visión nocturna los detectará. Se inventó para el ejército durante la segunda guerra mundial; se iría luego comercializando poco a poco. Ha avanzado por "generaciones" (como se denomina a las rigurosas normas de eficacia definidas por los Laboratorios de Visión Nocturna del Ejército de EE.UU. en Fort Belvoir). A finales de los años ochenta, las generaciones uno y dos cedieron el paso a los productos de la "generación 3" de hoy. Con la luna en cuarto (que da una intensidad luminosa de 0,01 lux), distinguen a una persona de un metro ochenta a más de quinientos metros. Los monoculares de calidad cuestan de dos mil a tres mil quinientos euros.

El corazón de cualquier aparato de visión nocturna es el tubo intensificador, que convierte en electrones los pocos fotones disponibles. Amplifica esa corriente y la convierte por último en imagen. Los intensificadores se degradan (los mejores brindan 10.000 horas de funcionamiento), pero son sustituibles. Funcionan mal, o quizá no funcionen, en presencia de humo, tormentas de arenas y niebla. Por ello, algunos soldados usan gafas infrarrojas, que captan los gradientes térmicos del enemigo o de un vehículo, o el gradiente térmico residual si acaban de pasar, para formar una imagen en blanco y negro. La nitidez de ésta es muy inferior a la que ofrecen los tubos de visión nocturna; si la temperatura de los objetos observados es uniforme, caso de un bache en una carretera, no aparecerá nada. Pero los dispositivos infrarrojos detectan objetos a kilómetros de distancia.

Por consiguiente, los fabricantes empiezan a producir prototipos de gafas "combinadas", que superponen una imagen infrarroja sobre la imagen en verde de un tubo intensificador. Puede que se conviertan en la próxima generación de aparatos de visión nocturna.



1. UN OBJETIVO MONOCULAR enfoca los escasísimos fotones disponibles en la oscuridad hacia un tubo intensificador, en cuyo interior reina el vacío. Allí, un fotocátodo (*1*) los convierte en electrones. Estos son acelerados por un voltaje en dirección a la placa microcanal (*2*), situada sólo a una distancia algo menor que una décima de milímetro, el grosor de dos hojas de papel. La placa multiplica los electrones y un segundo voltaje los acelera hacia la pantalla fosforescente (*3*), también a una distancia inferior a una décima de milímetro. Allí, los electrones se reconvierten en los fotones que llegan al ocular.

¿SABIA USTED QUE...?

➤ **UN RETO DE 6000 VOLTS:** El fotocátodo, la placa microcanal y la pantalla fosforescente del interior de un tubo intensificador están casi en contacto, pero los voltajes que aceleran los electrones que atraviesan los minúsculos intersticios existentes entre esas piezas varían entre 500 y 6000 volts. "La mayor dificultad técnica es mantener las placas paralelas y tan cercanas, y sin embargo tan limpias, que puedan aplicarse unas tensiones de esa magnitud sin que haya descargas disruptivas", según Tom Peck, de Visión Nocturna de Industrias ITT. Dos pilas AA pueden proporcionar durante 50 o 60 horas esos voltajes para unas corrientes de muy baja intensidad.

➤ **LUCES DE LA CIUDAD:** A los primeros aparatos de la generación tres les cuesta mantener las imágenes cuando la iluminación es muy variable, como en una ciudad. No serán muy útiles, pues, para los soldados que participen en operaciones militares en suelo urbano. Controlando la tensión por medio de puertas electrónicas y rebajando la corriente eléctrica dentro del intensificador, los fabricantes han aumentado el margen dinámico de los aparatos. Ahora un soldado puede divisar a un enemigo oculto en las sombras dentro de un callejón iluminado y mantenerlo localizado aunque los faros de un vehículo que pase alteren la escena.

➤ **NO ESTA A LA VENTA:** Los monoculares de visión nocturna, los gemelos de campaña, las gafas y cámaras especiales, las alzas de cañón y otros equipos están declarados de "exportación restringida" por el Departamento de Estado de EE.UU. Toda persona o empresa que venda o envíe al extranjero esos dispositivos sin autorización oficial podría enfrentarse a fuertes multas e incluso a años de cárcel.



2. LAS IMAGENES DE VISION NOCTURNA son verdes porque los luminóforos de los instrumentos emiten luz en las proximidades de la longitud de onda verde de 550 nanómetros, que es el "máximo fotópico" para el ojo humano.

3. LA PANTALLA FOSFORESCENTE emite una imagen visible cuando en ella inciden los electrones acelerados, de un modo muy parecido a lo que ocurre en un tubo de rayos catódicos. Una película impide la reflexión de los fotones hacia el fotocátodo, que causaría una realimentación destructiva.

4. EN LA PLACA MICROCANAL hay millones de minúsculos canales inclinados. Al entrar en ellos, los electrones rebotan en sus paredes. Cada choque produce dos o tres electrones más; su número se multiplica así varias veces. Una película impide que los iones positivos creados en algunos choques formen una corriente de retroceso hacia el fotocátodo, porque lo degradaría. Entre ambas caras de la placa hay un voltaje que impulsa a los electrones por los microcanales.

5. EN LAS CAPAS DEL FOTOCATODO, de sólo unas micras de espesor, inciden los fotones, cuya energía arranca electrones. El arseniuro de galio reacciona vivamente a las frecuencias de la luz emitida por el cielo nocturno, que caen sobre todo en el rojo y cerca del infrarrojo.

Ciencia romántica

Goethe

The Romantic Conception of Life. Science and Philosophy in the Age of Goethe, por Robert J. Richards. The University of Chicago Press; Chicago, 2002.

The Will to Create. Goethe's Philosophy of Nature, por Astride Orle Tantillo. University of Pittsburgh Press; Pittsburgh, 2002.

Carl von Linné: Nemesis Divina. Texto crítico y traducción de M. J. Petry. Kluwer Academic Publishers; Dordrecht, 2001. La Science Goethéenne des Vivants. De l'Histoire Naturelle à la Biologie Evolutionniste, por Jean-Michel Pouget. Peter Lang, Berna, 2001.

The Laboratory of Poetry. Chemistry and Poetics in the Work of Friedrich Schlegel, por Michel Chaouli. The Johns Hopkins University Press; Baltimore, 2002.

A bote pronto, todos asentiríamos a la afirmación de que entre ciencia y romanticismo existe *contradictio in terminis*. Nos lo pensaríamos dos veces si, volviendo la mirada a la biología del siglo XIX, la que desemboca en la teoría evolucionista de Darwin, descubriéramos que ésta, como ya reconociera el propio Haeckel, no es más que la coronación de la tesis de Goethe sobre los arquetipos. Pero quizá nos siga pareciendo escasa de rigor y sobrada de humo la divisa del Romanticismo: "El arte debe convertirse en ciencia y toda la ciencia en arte; poesía y filosofía deben ser una". La acuñó uno de sus máximos exponentes, Friedrich Schlegel (*The Romantic Conception of Life. Science and Philosophy in the Age of Goethe, The Laboratory of Poetry. Chemistry and Poetics in the Work of Friedrich Schlegel*).

¿Podemos o no hablar de ciencia romántica? Podemos, si la contextualizamos en su momento histórico. Y si trascendemos la visión alicorta de considerar en Goethe sólo al autor inmortal del *Fausto,* de *Las tribulaciones del joven Werther* o de las *Elegías romanas*. El mismo apreciaba su obra científica, que reputaba revolucionaria, muy por encima de su creación poética. Aunque, según expone con exhaustiva documentación Richards, en el Romanticismo la literatura, la filosofía, la investigación científica y la propia aventura personal son factores inseparables. Esa integración se expresa en su alejamiento del modelo mecanicista y la sustitución del mismo por otro organicista.

Cartesianos y newtonianos, con toda la pléyade racionalista que llega hasta el Kant de la *Crítica de la razón pura*, compartían una explicación del mundo, inerte o animado, basada en leyes físicas, ahormadas en parámetros matemáticos. Contra ese planteamiento se rebela la *Naturphilosophie*, en la que se encuadra la ciencia romántica. Propone ésta, por lo que respecta a la naturaleza viva, la existencia de tipos orgánicos fundamentales, arquetipos (*Archetypi, Urtypen, Hauptypen, Urbilden* y similares). En los textos introductorios de zoología todavía suele recogerse la división cuatripartita de las estructuras básicas: radiata (por ejemplo, estrellas de mar y medusas), articulata (insectos y crustáceos), mollusca (bivalvos y pulpos) y vertebrata (de los peces al hombre). Estos arquetipos, a su vez, se incluían en uno más fundamental: el plan animal general o vegetal general.

Para Kant, los arquetipos encarnaban un ideal que sólo existía realmente en la mente divina. Al científico competía abordarlos desde una perspectiva newtoniana, de lo que discrepaban los *Naturphilosophen*, quienes rechazaban cualquier planteamiento mecanicista siquiera fuera metodológico. Eso no empece que la filosofía kantiana de la biología ejerciera un influjo decisivo en el pensamiento de los románticos. Se expone en la *Crítica del juicio*, que versa sobre los juicios estéticos y teleológicos. Los segundos son juicios sobre el carácter finalista de la naturaleza orgánica. Kant concede carácter causal a la finalidad (*Zweck*). En el ejemplo clásico del edificio, éste constituye un fin porque es el producto de un concepto que existe en la mente del arquitecto. En el mundo vivo la finalidad ocupa un puesto preeminente, con la peculiaridad de que los mismos órganos son medios desde una perspectiva y fines desde otra. Todos, empero, operan y se ordenan hacia un fin último, el estado de perfección del organismo. Este horizonte determina la razón de ser y de actuar de cada órgano. Si tal concepto holístico lo aplicamos a un grupo animal extenso, tendremos el *arquetipo*, bajo cuya luz el biólogo establece el funcionamiento de un individuo, de una especie.

Teleología y holismo serán dos propiedades distintivas del pensamiento de Goethe, quien vino "al mundo en Frankfurt del Main el 28 de agosto de 1749, cuando el reloj daba las doce de mediodía" (comienzo de su *Dichtung und Wahrheit*). Nacido en el seno de una familia adinerada, recibió de su padre el amor a las lenguas clásicas y modernas. En 1765, se matriculó en la Universidad de Leipzig, para estudiar leyes, carrera que prosiguió en Estrasburgo (1770-71), donde coincidió con quien había de ser guía de su juventud, Johann Gottfried Herder, antiguo discípulo de Kant y ardiente defensor de la cultura alemana. Goethe fue llamado, en 1776, a la corte de Weimar. Nombrado superintendente de minas, contó para las labores de inspección con la ayuda de Johann Carl Wilhelm Voigt, alumno de Abraham Gottlob Werner. A esa época pertenecen sus primeras incursiones en el origen de las rocas, asunto que enfrentaba a neptunistas y vulcanistas.

Intentó, por su parte, una solución de compromiso: en un comienzo la Tierra se hallaba en un estado fluido; un fuego interno mantenía los minerales en solución; el granito, precipitado de un mar en ebullición, rodeó al núcleo formando cristales de múltiples capas. El enfriamiento del mar depositó otras especies rocosas que penetraron en el granito.

Por ese tiempo, Goethe conoció a Justus Christian Loder, profesor de medicina en Jena. Durante un año estuvo asistiendo a sus sesiones anatómicas. Ambos pertenecían a la misma logia masónica, avanzando juntos varios grados. Goethe, que ya empezaba a dominar el dibujo, se interesó por la anatomía comparada. Se centró en un tema un tanto obscuro: averiguar si los seres humanos tenían, igual que otros vertebrados, el hueso intermaxilar (*Zwischenkiefer* en alemán). Negar su presencia era negar la unidad de la naturaleza. No lo veían así Pieter Camper, Samuel Thomas von Soemmerring, Johann Friedrich Blumenbach y otras celebridades de la medicina contemporánea, quienes consideraban, en efecto, tal ausencia un signo de la separación radical que mediaba entre el hombre y el resto de los animales. Goethe defendió abiertamente, en 1784, la presencia del intermaxilar en *Dem Menschen wie den Tieren ist ein Zwischenknochen der obern Kinnlade zuzuschreiben* ("Lo mismo que a los animales, se atribuye al hombre un hueso intermaxilar de la mandíbula superior").Había llegado a tal descubrimiento por el camino de la embriología.

De la craneología comparada (que le llevó a inferir que los huesos del cráneo eran vértebras modificadas) pasó a estudiar las relaciones entre animales y plantas. Apoyado en sus observaciones microscópicas, expuso su convencimiento de que los animales más elementales procedían, por metamorfosis, de las plantas. (Era en 1786, en vísperas de su partida a Italia.) Se había iniciado en el mundo vegetal de la mano de Linneo, en quien veía un precursor de la unidad de plan, si bien de las ideas del sueco sobre un orden regulador de todo, lo biológico y lo moral, sólo conocía retazos (*Carl von Linné: Nemesis Divina*). En las últimas ediciones de su *Systema Naturae* Linneo introduce, como una extensión de su obra científica, su concepción de la ley natural. Petry ha agavillado lo que a primera vista no es más que un conjunto de apuntes sin vertebrar, expuestos en forma de casos y en un estilo sucinto, casi aforístico. Lo ofrece en espléndida edición traducida al inglés tras colacionar el principal texto autógrafo de la *Nemesis,* que se encuentra en la biblioteca de la Universidad de Uppsala, y material manuscrito conservado en la Sociedad Linneana de Londres y en el Instituto Carolino de Estocolmo.

Para Linneo, la ley natural es la ley del orden de la naturaleza. También Newton veía leyes inmutables tras el movimiento de los cielos o la caída de los gra-

"Hombre y mujer contemplando la Luna", *ca. 1824. Galería Nacional, Museo estatal de Berlín*



ves. Pero eso parecía afectar sólo a la materia inerte, cuyo comportamiento podía determinarse a través de ecuaciones rigurosas, matemáticas. Más difícil resultaba descubrir el orden divino en los organismos, donde todo fluía, cambiaba de forma, se desintegraba y se generaba. Aunque sin demostrarlo, Linneo afirmaba la existencia de un orden inexorable en la naturaleza, manifestación de la ley inmutable de Dios (*lex immutabilis Dei*). Bajo esa ley incluía al hombre, con su vertiente moral. Un caso paradigmático puede ser el del incesto. A su enunciado general —"El matrimonio o cópula entre consanguíneos debe prohibirse"— sigue un desarrollo donde, con intención, se infieren de la genética (dejemos aparte el carácter erróneo de alguna afirmación) conclusiones morales, porque para Linneo todo es un *continuum*: "La descendencia deriva su tuétano de la madre, del padre su cuerpo. Las enfermedades hereditarias que se perpetuarían con el matrimonio entre consanguíneos, se debilitan y erradican con el matrimonio entre alejados. Los criadores de caballos no son inteligentes cuando cruzan animales nacidos del mismo macho. Los extraños entre sí, al cruzarse, originan variedad y diversidad, de cuerpo y alma. A la naturaleza le gusta la variedad. Por tanto, las buenas cualidades que han permanecido confinados en una familia se reparten así entre muchas. Es cierto que algunos matrimonios entre consanguíneos raramente son un éxito. El obispo luterano Mennander se casó con su prima para demostrar que era incierto que dos consanguíneos no podían formar un buen matrimonio. Murió ella en su primer posparto".

Pese a su aprecio de la obra linneana no discurría por ese cauce la idea de unidad que acariciaba Goethe, y que se le hizo patente, con la fuerza de una visión, durante su estancia en Italia. Cruzó los Alpes en septiembre de 1786. Pasó por Verona y se detuvo en Venecia dos semanas. Siguió a Florencia. De allí a Roma, donde se quedó cuatro meses. Bajó a Nápoles. Se embarcó rumbo a Sicilia. En Palermo, el 6 de abril de 1787 acudió al jardín botánico para disfrutar con la lectura de la *Odisea*. Como una inspiración le vino a la mente que, bajo tamaña eclosión de especies, debía existir una *Urpflanze*, un arquetipo de todas las plantas. Fue entonces cuando propuso la hipótesis de que "todo es hoja y, en razón de su simplicidad, cabe la mayor multiplicidad posible". La hoja podría transformarse en otros órganos. "A la hoja que sólo absorbe nutrientes bajo la tierra la llamamos raíz; a la hoja que se forma con esos nutrientes la llamamos bulbo." Pero la hoja debe entenderse aquí simbólicamente, en cuanto fuerza dinámica unitaria que subyace bajo las múltiples transformaciones a las que da base. Goethe volvió a Weimar en junio de 1788. Entre 1788 y 1790 no sólo compuso sus *Römische Elegien*, sino que asentó dos líneas de investigación que le darían fama: sus investigaciones en óptica (*Beiträge zur Optik*, 1791-92) y en morfología. Aunque su obra principal sobre los colores (*Zur Farbenlehre*) no apareció hasta 1810, Goethe se lanzó a un ataque estéril contra la demostración newtoniana del carácter compuesto de la luz blanca. Más éxito tuvo su apartado sobre los colores fisiológicos (que dependen de la condición del ojo más que de la iluminación).

En 1790 apareció *Die Metamorphose der Pflanzen* (La metamorfosis de las plantas). Hermann Helmholtz, por otra parte sumamente crítico de la ciencia goethiana, reconocería que la exposición de su morfología determinó el planteamiento de la biología de la primera mitad del siglo XIX. A través de la observación empírica y la comparación, se propuso acotar el contenido conceptual de la *Urpflanze*. Lo asoció a un poder activo que daría origen a un número ilimitado de variedades vegetales. Debía existir una ley que gobernase las estructuras de todas las plantas. Su *Metamorfosis* sigue el desarrollo de las plantas a través de diversos estadios de transición. Los primeros órganos epigeos en aparecer eran las hojas seminales, los cotiledones. Mientras que las hojas en su crecimiento por el tallo forman una espiral, los cotiledones constituyen inicialmente un par opuesto, adheridos a un nodo. Pronto del nodo emerge un botón, una plúmula. Se irá formando el tallo y aparecerán gradualmente las hojas. Se alcanza otro estadio de desarrollo cuando las hojas apicales del tallo, los sépalos, dan lugar a una copa, o cáliz. El cáliz se contrae y así las hojas coloreadas, los pétalos, se juntan para producir la flor. Durante el proceso de la floración, varios pétalos se contraen en estambres y pistilo. Con la polinización, el fruto y la semilla que encierra se forman en una cápsula, cuyo aspecto foliar es manifiesto.

Goethe describió los diversos estadios del desarrollo de la planta como una expresión de las fuerzas universales de expansión y contracción (*La Science Goe-*



Friedrich Schlegel, alrededor de 1800. Artista desconocido

théenne des Vivants. De l'Histoire Naturelle à la Biologie Evolutionniste). Lo que nos lleva a su concepción de la filosofía natural (*The Will to Create. Goethe's Philosophy of Nature*). Aunque en línea con los estudios precedentes de Margrit Wyder, Dorothea Michaela Noë-Rumberg y Günter Schnitzler y Gottfried Schramm, Tantillo hace hincapié, sin embargo, en la condición dinámica de la naturaleza, según Goethe. Se centra en los principios fundantes de su teoría de la naturaleza, que se pretenden revolucionarios. Revolucionarios porque los opone a Descartes y a Newton. En opinión de Goethe, la ciencia de la Ilustración había absorbido tanto a las mentes, que les había robado la capacidad y, sobre todo, la libertad de perseguir ideas científicas fuera del canon aceptado. De un modo plástico, quería sustituir las construcciones matemáticas, deductivas y analíticas al uso por nuevos constructos fenomenológicos, capaces de dar cuenta de la actividad fluida y a veces impredictible de la naturaleza.



Johann Wolfgang von Goethe (1749-1832)

Las objeciones de Goethe a la metodología de la ciencia moderna no se ceñían al estilo formal, matemático. Cuestionaba también la idoneidad del lenguaje empleado. La expresión escrita tenía que "entrar por los ojos", única forma de aprehender todos los entresijos de la naturaleza. El desarrollo "teatral" de un experimento científico —con esa imagen lo entiende él— debe quedar reflejado en la exposición literaria del mismo. Para ello debe ésta ayudarse del dibujo y la pintura. Otro romántico, Schlegel, seguirá el camino inverso e impregnará su obra literaria de términos y metáforas tomadas de la química (*The Laboratory of Poetry. Chemistry and Poetics in the Work of Friedrich Schlegel*). La química del comienzo del siglo XIX era una ciencia en plena transformación, de la alquimia a la química. Para los románticos, la química mediaba entre la visión mecanicista y la visión organicista del mundo. Más de un decenio antes de que el propio Goethe se sintiera atraído por las ideas químicas en sus *Wahlverwandtschaften* ("Afinidades electivas"), publicadas en 1808, Schlegel se apoya en la química para renovar la estética. Schlegel no desarrolló ningún modelo químico, por eso no aparece en la historia de la ciencia. Pero habla por doquier de "condensación", "síntesis y análisis", "oxidación", "saturación", "nitrógeno" y, sobre todo, de fuerzas en eterna división y mezcla. Aunque en puridad, parangonado con Kant, Novalis o Goethe, Schlegel no pasa de ser un diletante.

Para explicar el comportamiento humano Goethe presenta, en efecto, una teoría química en *Wahlverwandtschaften*. El hombre, viene a señalar, no se halla por encima de la naturaleza, sino que se insiere en ella. Pero, ¿cuáles son los principios que aúnan a la naturaleza entera? Se reducen a cuatro fundamentales: polaridad, intensificación (*Steigerung*), composición y competición, enlazadas por el hilo conductor de la fuerza creadora. Así, para Goethe, la polaridad constituye uno de los principales impulsos de los procesos naturales, del orden que sea. Se trata del lenguaje propio de la naturaleza. El universo entero consta de entidades opuestas que se disgregan y vuelven a juntarse engendrando nuevas entidades, en un estado de flujo constante.

En razón del principio de intensificación, cada una de las partes de la naturaleza pugna por ascender a un grado de mayor complejidad y perfección. Es la fuerza operativa que hay detrás de la metamorfosis de una planta en desarrollo. Recurre a este principio para rechazar las ideas botánicas de Linneo y de Caspar Friedrich Wolff. De él se vale para respaldar el concepto de impulso formador (*Bildungstrieb*) propuesto por Johann Friedrich Blumenbach. Si Aristóteles limitaba el impulso formador a la procreación, la *Steigerung* goethiana permite explicar una naturaleza, inerte o no, que se supera a sí misma.

El principio de compensación atiende a las diferencias intra e interespecíficas. Tomando uno de sus ejemplos, la serpiente "satisface" con su cuerpo alargado la falta de extremidades. Por último, relaciona el principio de competencia con el género y la reproducción, extendida esta última para cubrir el desarrollo. La pugna competitiva se manifiesta con nitidez en las tendencias masculinas y femeninas del mundo vegetal. Dentro de una misma planta unas veces puede predominar lo masculino y otras lo femenino.

—LUIS ALONSO

Hitos de la era espacial

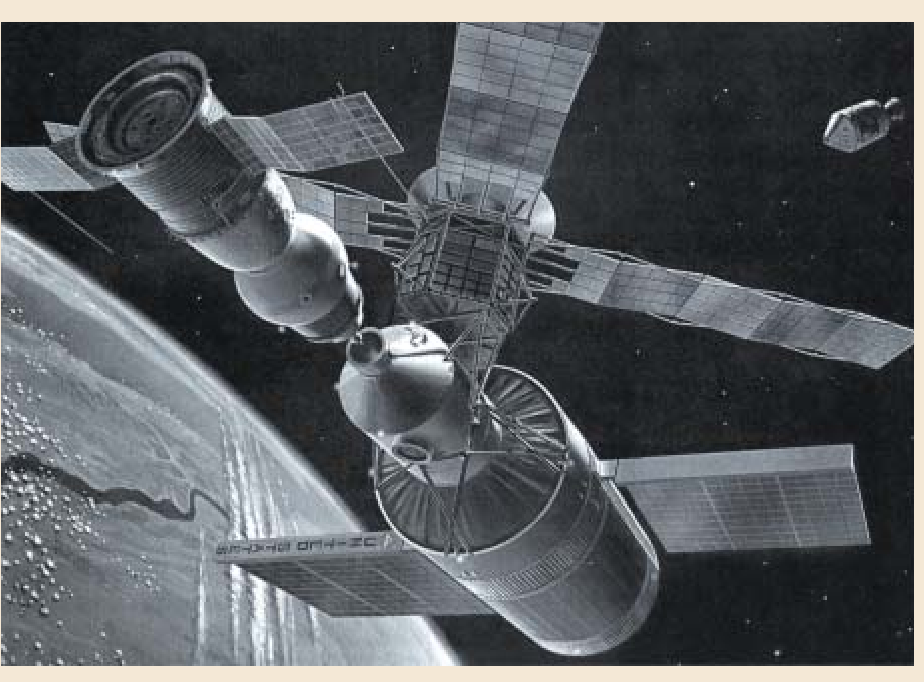
TRACKING APOLLO TO THE MOON, por Hamish Lindsay. Springer Verlag; Berlín, 2001. **SKYLAB**, por David J. Shayler. Springer; Berlín, 2001. **THE ROCKET MEN**, por Rex Hall y David J. Shayler; Springer; Berlín, 2001.

Edad Antigua, edad Media, edad Moderna, edad Contemporánea. Tal es la división clásica de la historia como se nos ha enseñado en la escuela a los occidentales, y según nuestra propia visión. El tránsito de una a otra nos lo ofrecen marcado por un hecho que, por sus características, presentaba los signos de hito; así, el paso de la edad Media a la Contemporánea, a la caída de Constantinopla en manos otomanas, o el paso de la edad Moderna a la Contemporánea, con la Revolución Francesa. Pero en los pasos de una edad a otra, la concreción de una fecha, la de la entrega del puente de Gálata o la de la toma de la Bastilla, no es sino una referencia que, en cada caso, pretende dar cuenta del remate de un proceso que se inició años, algunas generaciones, antes.

Superponiéndolas a esta división clásica de la historia se habla de eras: períodos caracterizados por una gran innovación. ¿Cuál sería el hito que señalaría el inicio de la era espacial? (Esta expresión, "era espacial", parece destinada a hacer fortuna, acaso mucho más que la de "era atómica", tan en boga hace cuarenta o cincuenta años.) Quizá sería propio responder que el inicio de la era espacial habría que buscarlo en el momento en que por primera vez alguien, hace miles de años, miró al firmamento con afán inquisitivo. Ese momento no lo conocemos. Pero si hemos de buscar una referencia con la pretensión de sentar un hito como si se tratara de marcar el paso de una edad histórica a otra edad histórica, muchos se inclinarían a señalar la noche del 20 al 21 de julio de 1969 en que Neil Armstrong, jefe de la misión Apolo 11, fue el primer humano que pisó suelo lunar.

Tracking Apollo... narra cómo pudo materializarse ese instante. El autor, Hamish Lindsay, ofrece un testimonio de primera mano, con material en gran parte recogido por él mismo a lo largo de los años que, a partir de 1963, sirvió en la NASA. Tras un interesante resumen de antecedentes históricos, procede a explicarnos las distintas fases que culminaron en la noche del desembarco en la Luna. El proyecto Mercury y el programa Géminis se destinaron, sólo a ello y desde un principio, a abrir el camino hacia la Luna. El primero tenía como propósitos esenciales colocar un hombre en órbita terrestre, comprobar que un hombre podría desenvolverse en ese entorno y recuperar intactos al hombre y al ingenio que lo había transportado. Por su parte, el programa Géminis pretendía conocer de qué modo podría actuar eficazmente el hombre en el espacio ultraterrestre. Más que nada, estudiar los efectos de la ingravidez sobre el organismo en misiones largas, para así determinar la capacidad de un ser humano para trabajar en el exterior, en un vacío sin referentes inmediatos. Importante fue también el objetivo de estudiar las maniobras de lanzamiento de un ingenio que acuda a una cita con otro y el acoplamiento de ambos, lo cual es esencial para operar en el espacio. Con la misión Géminis 12, entre el 11 y el 15 de noviembre de 1966, se cierra prácticamente el programa Géminis. El proyecto Apolo fue anunciado el 28 de julio de 1960 con el objetivo declarado de llevar a cabo un aterrizaje tripulado en la superficie lunar, para lo cual se pondría en órbita, alrededor de la Luna, un satélite artificial desde el que se desprendería un vehículo especial que aterrizaría en nuestro satélite y cuya tripulación realizaría determinadas observaciones, recogería muestras, dejaría una instrumentación que transmitiría información hacia la Tierra y, finalmente, regresaría al satélite y a la Tierra. Para ello era necesario desarrollar todo un sistema de equipos e instalaciones. La misión Apolo I programada para el 27 de enero de 1967 acabó en desastre al incendiarse los tanques de oxígeno y morir los tres tripulantes. El proyecto prosiguió para culminar en la misión Apolo 11. A esta sucedieron seis misiones más, hasta la Apolo 17, que entre el 7 y el 19 de diciembre de 1972 llevó a los últimos, hasta ahora, astronautas a la Luna. Se cierra así el proyecto Apolo, para dar paso al programa Skylab. El autor ofrece una narración ante todo muy entretenida, pero también completa y detallada de la historia de los primeros vuelos espaciales tripulados norteamericanos. Se trata esencialmente de un libro de historia, escrito por alguien que intervino directamente en ella y no sólo como testigo. Ello hace que determinados fragmentos, como por ejemplo los que narran los pormenores del primer alunizaje, resulten especialmente vívidos y directos. Puede que esta obra sea una referencia obligada para quienes se interesen en la historia de la era espacial, no sólo respecto a la evolución de los hechos escuetos sino también a su relación con los aspectos técnicos y al contexto de ambos, durante lo que se llamó "carrera espacial" en los años crudos de la guerra fría. La obra termina con un breve capítulo dedicado a la misión Apolo-Soyuz, primera cita en el espacio entre rusos y norteamericanos a mediados de julio de 1975.

Se nos narra en *Skylab* la historia del programa de ese nombre, que en EE.UU. sucedió al proyecto Apolo. Su autor, David J. Shayler, ha recopilado la mayor cantidad de información reunida hasta ahora sobre el tema bajo un solo encabezamiento. El programa Skylab adquirió carta de naturaleza como conjunción de un número de proyectos e ideas que en



Recreación artística de un acoplamiento de la sonda Soyuz al Laboratorio Espacial

años anteriores se habían propuesto sin que, esencialmente por falta de fondos, salieran adelante. Para su puesta en ejecución se aprovecharon los equipos e instalaciones del proyecto Apolo, pero con una finalidad más amplia. Hasta entonces, la actividad espacial apenas si había rebasado la frontera de la ingeniería, con la atención centrada en cuestiones tales como las condiciones de alojamiento de la tripulación, ensayo de procedimientos, equipos e instrumentación, y ejecución de maniobras de aproximación, acoplamiento y alunizaje.

La idea de colocar un laboratorio en el espacio, el Skylab, estaba alimentada por la intención de recoger datos de interés, fueran científicos, médicos o técnicos, con vistas a asegurar una presencia permanente y activa del hombre en el espacio. Entre mayo de 1973 y febrero de 1974 hubo tres visitas al laboratorio puesto en órbita el 15 de mayo de 1973. A través de sus más de 350 páginas, el libro sigue la pista, desde los antecedentes, inicios y preparativos, de todos los hechos que jalonan la historia del programa Skylab, considerado éste como paso adelante decisivo para los avances posteriores y que actualmente se plasman en la Estación Espacial Internacional. Por ello, no se limita el autor a la narración de hechos y a la explicación de cuestiones técnicas, expone además los conocimientos y experiencias adquiridos en cada misión, lo que al fin y al cabo eran la justificación última del programa. Es, pues, éste también un libro de historia y quizá también de referencia para los que estén empezando a interesarse por la historia de la era espacial, y que por otra parte puede atraer a cualquier clase de lectores.

Aunque admitamos que el hito iniciador de la era espacial haya que colocarlo en el momento del alunizaje de la misión Apolo 11, es innegable que otros pueden sustentar la opinión de que ese hito debería colocarse en el 12 de abril de 1961, cuando Yuri Gagarin entró en órbita terrestre tripulando un satélite artificial de la clase Vostok, convirtiéndose así en el primer hombre que penetraba en el espacio ultraterrestre. Por razones que solemos considerar de sobra conocidas, los comienzos de la aventura espacial rusa, o más exactamente soviética, no tuvo entre nosotros la inmediatez de su contrapartida americana. En aquellos años, los detalles e incidencias de la carrera soviética hacia el espacio permanecían envueltos en el secreto. Capítulo de antecedentes históricos aparte, *The Rocket Men*, de Rex Hall y David J. Shayler, cubre el desarrollo y los logros, en la URSS, en el campo de los vuelos espaciales durante los dos primeros lustros de su historia activa, situando el comienzo de ésta a mediados de los años cincuenta, con el lanzamiento del primer satélite de la clase Sputnik el 4 de octubre de 1957, hecho que realmente sorprendió a todo el mundo. *The Rocket Men* trata de los satélites Sputnik y Vostok, de las naves Vosjod, de los vehículos de lanzamiento y del cosmódromo de Tyuratam, y también de las personas que intervinieron en aquellos proyectos, y de cómo la combinación de esos factores hizo posible que la entonces Unión Soviética consiguiera adelantarse en objetivos tales como satélites en órbita y el primer vuelo espacial humano, tiempos de permanencia en el espacio, lanzamientos pilotados, vuelos con varios tripulantes y paseos espaciales.

Hoy, al cabo de casi cuarenta años, todo ello nos parecen temas de noticias casi de rutina, pero en su época eran noticias merecidamente de primera plana. Escrito con el mismo ánimo y las mismas pautas que *Skylab*, con el que comparte autor, el libro cierra su narración con la misión Cosmos 110, en la que se empleó una nave Vosjod con dos perros a bordo para estudios biomédicos y que fue lanzada en febrero de 1966. Para la redacción y las ilustraciones se emplearon fuentes occidentales especializadas en los programas soviéticos, así como información procedente de Rusia, ya disponible desde 1991, y de material desclasificado de la inteligencia estadounidense. Todo ello hace de este libro una obra de la misma altura e interés que los otros dos objeto de esta reseña.

—José Vilardell

AVENTURAS PROBLEMÁTICAS

Dennis E. Shasha

Pinchazos en corto

Se encarga usted del servicio de transmisiones confidenciales de un departamento diplomático. Hay espías dispuestos a interceptar los mensajes. Para pinchar radioenlaces de microondas, los fisgones han de alinearse con las antenas emisora y receptora; ello los pone en peligro de ser detectados y detenidos, y cabe suponer que, si lo intentan, lo harán solamente una vez, y durante no más de 10 minutos. Claro está que no sabemos cuáles son los intervalos de 10 minutos que van a elegir.

Tiene usted que enviar siete mensajes, alfabéticamente ordenados, en el mínimo tiempo posible. Las duraciones correspondientes son:

A: 2 minutos
B: 3 minutos
C: 4 minutos
D: 5 minutos
E: 6 minutos
F: 7 minutos
G: 8 minutos

Puede tolerarse que los espías intercepten hasta tres de estos comunicados. Un mensaje se considera interceptado si es "pinchado" de principio a fin; los pinchados en parte, de nada le sirven al fisgón. Cada uno de nuestros mensajes ha de ser enviado sin interrupción; por otra parte, podemos enviar "en paralelo" tantos como queramos, así como empezar a enviar uno o varios mensajes mientras se están enviando otros.

Como ejercicio preliminar, supongamos aceptable que los espías logren examinar sólo un mensaje completo. ¿En qué momento deberíamos enviar cada mensaje para hacer mínimo el tiempo total de transmisión? En la figura podemos ver una solución en 36 minutos.

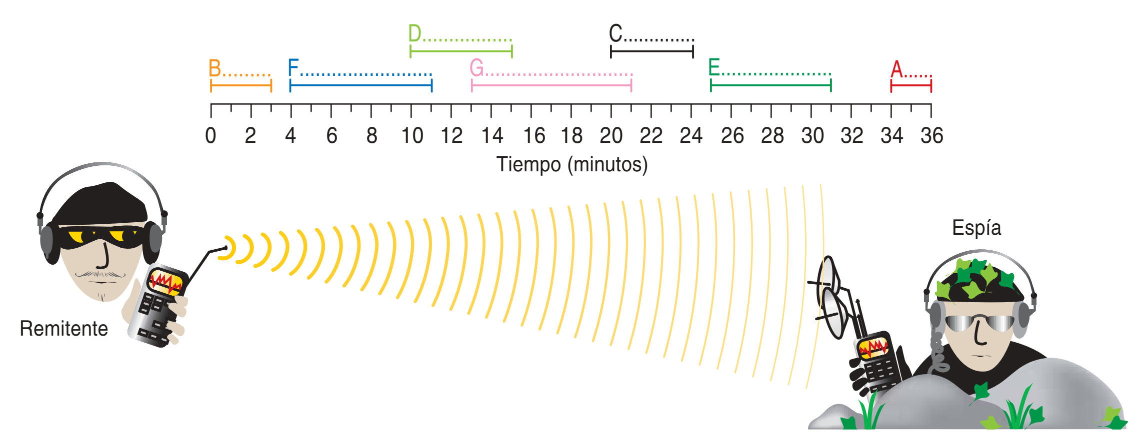
Dado que en el problema principal es permisible que se pinchen 3 mensajes, se podrá reducir el tiempo de manera considerable. ¿Logrará usted dar con la forma de enviar los siete mensajes en no más de 15 minutos? Y si además de estos siete mensajes tuviera que enviar otros tres mensajes de cuatro minutos, sin que fueran más de 3 de esos 10 los pinchados en su totalidad, ¿conseguiría hacerlo en 20 minutos o menos?

Solución del problema del mes pasado:

Al menos uno, y a lo sumo tres, contenedores viajan desde el puerto A hasta el puerto D. El país E recibe a lo sumo dos contenedores de A, y posiblemente, ninguno. En el segundo problema, D recibe a lo sumo dos contenedores llegados de A, y posiblemente, ninguno. Los números mínimo y máximo procedentes de A que llegan a E no varían.

Solución en la Red

Puede verse una explicación más completa de la solución del problema del mes pasado, así como la solución del problema de este mes, visitando www.investigacionyciencia.es.

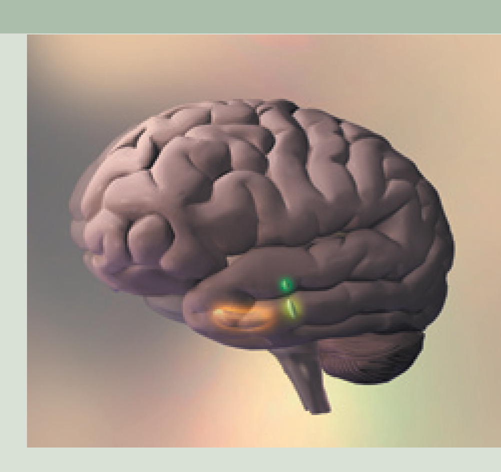


CORNELIA BLIK

UNA SOLUCION PRELIMINAR
Para enviar todos los mensajes en 36 minutos sin que más de uno pueda ser interceptado, se transmite B en el minuto 0, seguido por F en el 4, D en el 10, G en el 13, C en el 20, E en el 25 y A en el 34. Ningún intervalo de 10 minutos puede contener la totalidad de dos mensajes. Tal vez logre usted una solución mejor.

...DE NOVIEMBRE 2003



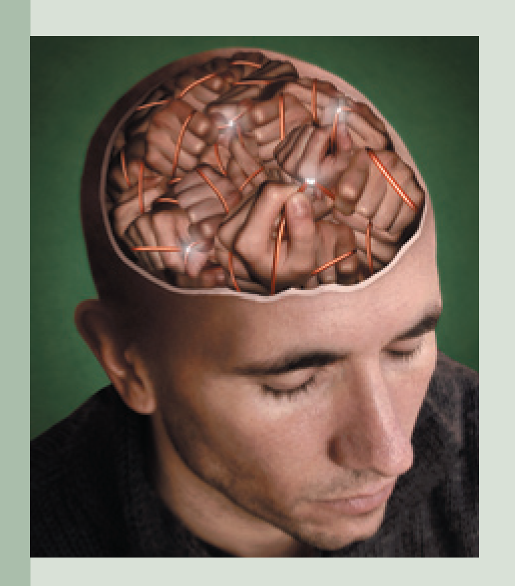


REGENERACION CEREBRAL, por Fred H. Gage

¿Puede repararse el cerebro dañado? Tal vez pudiera lograrse por el mismo camino que nos permita llegar a aumentar la potencia de un cerebro sano.

FARMACOS DE LA INTELIGENCIA, por Stephen S. Hall

Todavía no se expenden en las farmacias los compuestos que refuerzan la memoria u otras capacidades cognitivas y ya han suscitado un interesante debate social sobre su empleo.



ESTIMULACION CEREBRAL, por Mark S. George

La activación de los circuitos cerebrales mediante impulsos de campo magnético puede suavizar las depresiones, avivar las facultades cognoscitivas e incluso combatir la fatiga.

LECTORES DE LA MENTE, por Philip Ross

Equipos de escáner cerebral podrían dentro de poco discernir pensamientos rudimentarios y discriminar entre hechos y ficciones.

MUTABILIDAD CEREBRAL, por Marguerite Holloway

Al parecer, el cerebro puede mejorar de forma insospechada por medio de ejercicios mentales y físicos específicos.

CONTROL DEL ESTRES, por Robert Sapolsky

El avance en el conocimiento de los procesos cerebrales del estrés indica la conveniencia de ir dejando de lado los ansiolíticos químicos.



LA DIAGNOSIS DE LOS TRASTORNOS PSIQUICOS, por Steven E. Hyman

Las enfermedades psiquiátricas son a menudo difíciles de reconocer, pero algún día podrán detectarse mejor mediante el test genético y las técnicas de formación de imágenes.

LAS CERAMICAS ANTIGUAS DEL OCCIDENTE DE MEXICO, por Robert B. Pickering y Ephraim Cuevas

Los insectos necrófagos y las máculas minerales sirven de ayuda a un antropólogo forense y a un químico para determinar la autenticidad de figurillas elaboradas hace 2000 años.

¿ES LO MEJOR MEJORAR?, por Arthur L. Caplan

El punto de vista ético de un filósofo que aboga por la potenciación cerebral.

EVOLUCION MOLECULAR DIRIGIDA, por Miguel Alcalde

Al recrear en el laboratorio los procesos claves de la evolución —mutación, recombinación y selección—, se están preparando moléculas de nuevo cuño. Se condensan en meses o semanas los millones de años en que se mueve la escala temporal de la evolución.

